Open Source Intelligence Techniques

Resources for Searching and Analyzing Online Information

Sixth Edition



Michael Bazzell

Open Source Intelligence Techniques

Resources for Searching and Analyzing Online Information

Sixth Edition

Michael Bazzell

OPEN SOURCE INTELLIGENCE TECHNIQUES:
RESOURCES FOR SEARCHING AND ANALYZING ONLINE INFORMATION
Sixth Edition

Copyright © 2018 by Michael Bazzell

All rights reserved. No part of this book may be reproduced in any form or by any electronic or mechanical means including information storage and retrieval systems without permission in writing from the author.

Sixth Edition First Published: February, 2018

Project Editor: Y. Varallo

The information in this book is distributed on an "As Is" basis, without warranty. The author has taken great care in preparation of this book, but assumes no responsibility for errors or omissions. No liability is assumed for incidental or consequential damages in connection with or arising out of the use of the information or programs contained herein.

Rather than use a trademark symbol with every occurrence of a trademarked name, this book uses the names only in an editorial fashion and to the benefit of the trademark owner, with no intention of infringement of the trademark.

Due to the use of quotation marks to identify specific text to be used as search queries and data entry, the author has chosen to display the British rule of punctuation outside of quotes. This ensures that the quoted content is accurate for replication. To maintain consistency, this format is continued throughout the entire book.

Library of Congress Cataloging-in-Publication Data:
Application submitted

ISBN-13: 978-1984201577

ISBN-10: 1984201573

CONTENTS

About the Author I
Introduction II
CHAPTER 1: Prepare Your Computer 1
Antivirus 1
Malicious Software 3
System Cleaner 3
Firefox 5
Firefox Settings 6
Firefox Add-Ons 8
Script Blocking 13
Firefox Profile 23
JavaScript Bookmarklets 24
Chrome 25
Chrome Extensions 26
Tor Browser 29
Virtual Private Network 30
CHAPTER 2: Buscador Linux Virtual Machine 33
Virtual Machines 34
VirtualBox 35
Buscador Download 35
Buscador Installation 36
Snapshots 37
Buscador Browsers 40
Buscador Video Utilities 40
Buscador Applications 43
Bootable USB Devices 51
CHAPTER 3: Search Engines 57
Google 57
Google Operators 57
Google Search Tools 63
Google Custom Search Engines 65
Alerts 69
Bing 70
Bing Operators 70
Images 71
Archives 71
Translators 76
Groups 78
News 79
Newspapers 79

Tor Search Engines 85
International Search Engines 86
Yandex 87
Yandex Operators 88
Private Search Engines 89
FTP Search 89
IntelTechniques Search Engine Tool 93
CHAPTER 4: Social Networks: Facebook 95
Account Creation 95
Facebook Search: Standard 97
Facebook Search: People 98
Facebook Search: Posts 102
Facebook Search: User ID 104
Facebook Search: Friends 108
Facebook Search: Common Results 111
Facebook Search: ID Creation Date 114
Facebook Search: Businesses 114
Facebook Search: Events 116
Facebook Search: Live Video 118
IntelTechniques Facebook Search Tool 124
Facebook Search: Email 127
Facebook Search: Telephone Number 127
CHAPTER 5: Social Networks: Twitter 135
Twitter Search 135
Twitter Search Operators 138
Deleted Twitter Posts 141
Twitter Biographies 144
IntelTechniques Twitter Search Tool 145
TweetBeaver 147
Twitter Location Information 150
Tweet Deck 156
Twitter Analytics 157
CHAPTER 6: Social Networks: Others 165
Instagram 165
Instagram Private Accounts 167
IntelTechniques Instagram Search Tool 169
LinkedIn 171
IntelTechniques LinkedIn Search Tool 173
Contact Exploitation 175
Account Export Options 178
CHAPTER 7: Online Communities 183
Reddit 183
Deleted Content 184

Reddit Alternatives 189
Dating Websites 191
Forums 194
Online Prostitution 196
Craigslist 198
eBay 200
Amazon 202
IntelTechniques Communities Search Tool 204
CHAPTER 8: Email Addresses 207
Email Verification 207
Email Assumptions 208
Compromised Email Databases 211
Email Searching 212
IntelTechniques Email Search Tool 214
CHAPTER 9: User Names 217
User Name Search Engines 217
IntelTechniques User Name Search Tool 221
User Name Assumptions 221
CHAPTER 10: People Search Engines 227
People Search Engines 227
IntelTechniques Person Search Tool 233
People Search Combination 234
Resumes 236
Gift Registries 238
CHAPTER 11: Telephone Numbers 243
Carrier Identification 243
Caller ID Databases 244
Telephone Search Databases 250
Search Engines 253
IntelTechniques Telephone Search Tool 255
Voicemail Retrieval 258
Loyalty Cards 259
CHAPTER 12: Online Maps 261
Google Maps 261
Bing Maps 263
Additional Maps 263
Crowd-Sourced Street Views 264
Historic Imagery 266
IntelTechniques Maps Search Tool 267
Maps Manipulation 273
CHAPTER 13: Documents 275
Google Searching 275
Google Docs 276

Amazon Data 277
Presentation Repositories 278
IntelTechniques Documents Search Tool 279
Metadata 281
Rental Vehicle Records 282
Paste Sites 283
IntelTechniques Paste Sites Search Tool 283
CHAPTER 14: Photographs 285
Reverse Image Searches 285
IntelTechniques Reverse Image Search Tool 289
Twitter Images 291
Metadata 293
Image Manipulation 297
Image Forensics 298
CHAPTER 15: Videos 303
YouTube 303
YouTube Restrictions Bypass 304
IntelTechniques YouTube Search Tool 308
Reverse Video Searching 308
IntelTechniques Reverse Video Search Tool 312
Video Search Options 313
Video Search Archives 315
Video Closed Captions 316
Live Video Streams 317
Periscope 318
CHAPTER 16: Domain Names 321
Domain Registration 321
Domain Search Tools 322
Historical Registration Data 323
Visual Depictions 326
Website Monitoring 327
Domain Analytics 328
Robots.txt 330
Search Engine Marketing Tools 332
Shortened URLs 336
IntelTechniques Domain Search Tool 337
CHAPTER 17: IP Addresses 339
IP Address Location 339
IP Address Search 340
Wigle 342
Shodan 343
IntelTechniques IP Address Search Tool 345
IP Logging 346

CHAPTER 18: Government Records 353
County General Records 353
County Court Records 353
State Business Records 354
Date of Birth Records 355
Social Security Records 355
Vehicle Identification Number Search 356
Vehicle Registration Search 357
Campaign Contributions 358
Criminal Information 358
Voter Registration Records 361
Virtual Currency Records 362
CHAPTER 19: Software Applications 363
Video Utilities 364
Video Download 367
Video Metadata 369
Google Earth 370
Creepy 372
Exif Tool 373
HTTrack 374
4K Stogram 374
CamStudio 375
Lightshot Capture 376
SmartDeblur 377
FOCA 378
ExtractFace 380
SEO Spider 381
Domain Hosting View 381
IP Net Info 382
CCleaner 382
BleachBit 382
VeraCrypt 383
KeePassXC 385
Recuva 385
CHAPTER 20: Application Programming Interfaces (APIs) 387
Pipl 389
Full Contact 392
Flickr 396
Reverse Caller ID 397
Service Objects 398
TowerData 399
Have I Been Pwned 401
Hacked-Emails 402

CHAPTER 21: Android Emulation **405**
Genymotion 406
Genymotion Configuration 406
Google Apps Installation 409
Android Apps 412
Contact Exploitation 415
Virtual Device Cloning 416
Virtual Device Export 417
Additional Android Emulation Options 418
CHAPTER 22: Recon-ng **419**
Recon-ng Commands 419
Recon-ng Workspaces 421
Recon-ng Modules 422
Recon-ng Reports 424
CHAPTER 23: Radio Frequency Monitoring **431**
Hardware 431
Software to Find Radio 431
Public Frequencies 432
Wireless Monitors 435
Wireless Microphones 436
Online Databases 437
Online Streaming Frequencies 440
CHAPTER 24: OSINT Workflow Processes **443**
Email Addresses 445
User Names 446
Real Names 447
Telephone Numbers 448
Domain Names 449
Locations 450
CONCLUSION: **457**
INDEX: **458**

About the Author

Michael Bazzell

Michael Bazzell spent 18 years as a government computer crime investigator. During the majority of that time, he was assigned to the FBI's Cyber Crimes Task Force where he focused on open source intelligence, cyber-crime cases, and personal data removal methods. As an active investigator for multiple organizations, he has been involved in numerous high-tech criminal investigations including online child solicitation, child abduction, kidnapping, cold-case homicide, terrorist threats, and high-level computer intrusions. He has trained thousands of individuals in the use of his investigative techniques and privacy control strategies.

Michael currently works and resides in Washington, D.C. He also served as the technical advisor for the first season of the television hacker drama *Mr. Robot*. His books *Open Source Intelligence Techniques* and *Hiding from the Internet* have been best sellers in both the United States and Europe. They are used by several government agencies as training manuals for intelligence gathering and securing personal information.

INTRODUCTION

Sixth Edition

The previous (fifth) edition of this book was originally released in May of 2016. I assumed that it would be the final version, and stated in a few communication channels that it would be the last book I would write on the topic. In that book, I focused more on global techniques instead of specific resources in an attempt to get some extra mileage out of it. Since the first edition was released in 2012, I had been pushing out an updated version every year. The fifth edition seemed like the proper exit for the series. It was not because I was tired of online investigations. I may be more passionate now about collecting online evidence than I ever was before. I simply wanted to focus more energy toward other interests and opportunities, and I began spending a large amount of my time researching advanced privacy techniques.

In that down-time, I co-wrote *The Complete Privacy & Security Desk Reference*, and started a weekly podcast titled *The Complete Privacy & Security Podcast*. I also launched a new company dedicated to assisting other people in disappearing completely when bad situations arose. Whether conducting online data-mining removals for privacy; facilitating property purchases through the use of anonymous land trusts and LLCs for asset protection; or complete relocations to safe houses in the middle of the night for protection, it was a fascinating two years of research and execution.

In late 2017, I had the itch to begin writing about online research methods again. Earlier that year, I co-created a Linux virtual machine targeted toward research professionals that included numerous utilities never mentioned in my previous books. This pre-configured operating system gained a lot of public interest and we continue to update it twice yearly. Over the past two years, I updated my online research tools every month in order to continue to provide functional resources. I kept a running log of all of the changes that might need more explanation. In early 2018, I started documenting all of this, plus some of my favorite new Linux tools, in written form with anticipation of creating a supplement to the fifth edition of this book. Within a couple of weeks, I realized that the entire book should be re-written and released as a new edition. I have always self-imposed a "rule" in reference to my book revisions. The potential release must include at least 25% brand new material, 25% updated content, and 25% untouched stable and beneficial techniques. I believe that this sixth edition meets this criteria.

Keeping a book up to date about ways to access information on the internet is a difficult task. Websites are constantly changing or disappearing, and the techniques for collecting all possible public information from them are affected. While the fifth edition of this book is still highly applicable, a lot has changed over the past two years. Much of this book contains new techniques

that were previously not available. The Facebook Graph search options continue to grow considerably. I have also created several new online search tools to help with the investigative process. While Twitter and Instagram took away a few features, there is an abundance of new techniques available to all of us. Finally, a surge of Python tools has bombarded us with new capabilities never available before. It is a very exciting time for internet investigations.

The first chapter helps you properly configure your online investigation computer. It briefly discusses proper security protocols and free software. Great emphasis is placed on proper use of secure web browsers. A major change since the previous edition was the launch of Firefox version 57. In this update, all legacy add-ons were eliminated. If the add-ons were not upgraded to Firefox's new requirements, the tools no longer work. We lost some great resources, but this chapter will outline some new benefits.

A brand-new chapter explains the importance of virtual machines and instructs you on making your own or using a pre-configured option called Buscador. This virtual machine, co-created by David Westcott and myself, takes away the technical difficulties of installing custom Python applications, and leaves the user with a point-and-click environment ready for any type of investigation. Users of any skill level can now take advantage of Linux-based applications once restricted to those that understood programming and terminal prompts. With proper use of this system, you will no longer need to worry about viruses or malware. Dozens of applications, all included in Buscador, are explained in great detail in Chapter Two.

The remaining chapters are structured a bit differently from previous editions. Instead of trying to combine related topics into a single chapter, such as "Telephone Numbers & Addresses" or "Domains & IP Addresses", each category now has its own chapter. This allowed me to really delve into each topic and isolate the various techniques.

Fortunately, knowing methods for accessing data on one website often carries over nicely to other websites. This entire sixth edition was accurate as of February 2018. If, or more likely when, you find techniques that no longer work, use the overall lessons from the entire book to push through the changes and locate your content. Once you develop an understanding of the data, you will be ready to adapt with it. As always, I will publish updates to my online blog and free newsletter.

I will also post new video tutorials for the members of my online training program. You can access all of this, including my current investigation tools and links, on my website located at **IntelTechniques.com**. More importantly, please consider joining my free online forum at that address. This is where you will hear about all of the amazing OSINT techniques and methods that are being discovered every day from some of the brightest minds in online research. There are currently over 4,000 registered users, some of whom are active daily.

Open Source Intelligence (OSINT)

Open Source Intelligence, often referred to as OSINT, can mean many things to many people. Officially, it is defined as any intelligence produced from publicly available information that is collected, exploited, and disseminated in a timely manner to an appropriate audience for the purpose of addressing a specific intelligence requirement. For the CIA, it may mean information obtained from foreign news broadcasts. For an attorney, it may mean data obtained from official government documents that are available to the public. For most people, it is publicly available content obtained from the internet.

What is this book?

Overall, this book includes several hundred sources of free and open data which could identify personal information about anyone. All of the resources are 100% free and open to the public. Each resource is explained, and any creative search techniques involving the resource are detailed. When applicable, actual case examples are provided to demonstrate the possibilities within the methods. The book can be read in any order and referenced when a specific need arises. It is a guidebook of techniques that I have found successful in my investigations.

Locating this free online information is not the final step of OSINT analysis. Appropriate collection methods will be detailed and referenced. Whether the data you obtain is for an investigation, a background check, or identifying problem employees, you must document all of your findings. You cannot rely on the information being available online forever. A website may shut down or the data may be removed. You must preserve anything of interest when you find it. The free software solutions presented here will help you with that.

OSINT search techniques do not apply only to websites. There are many free programs that automate the search and collection of data. These programs, as well as application programming interfaces, will be explained to assist the advanced investigator of open source intelligence.

In summary, this book is to serve as a reference guide to assist you with conducting more accurate and efficient searches of open source intelligence.

What the book is not…

This is not a debate about the ethics or politics of online reconnaissance for personal information. It is not a historical look at OSINT or a discussion of administrative policy. There are better books that tackle these subjects. Furthermore, it is not a how-to guide for criminals to steal your identity. Nothing in this book discusses illegal methods of obtaining information.

Book Audience

When I first considered documenting my OSINT techniques, the plan was to post them on my website in a private area for my co-workers. This documentation quickly turned into over 250 pages of content including screen shots. It had grown too big to place on my site in a manner that was easy to digest. I changed course and began putting together this book as a manual to accompany my multiple-day training sessions. I now hope that a wider investigation community can gain something from these techniques.

Many readers are in some form of law enforcement. Police officers can use these techniques to help locate missing children or investigate human trafficking. Intelligence analysts can apply these methods to a large part of their daily work as they tackle social media posts. Detectives can use the search techniques to re-investigate cases that have gone unsolved.

I now offer my online and live OSINT training to the private sector, especially global security divisions of large corporations. This book can help these teams locate more concise and appropriate information relative to their companies. These methods have been proven successful for employees that monitor any type of threat to their company, from physical violence to counterfeit products. I encourage the use of these techniques to institutions that are responsible for finding and eliminating "bad apples". This may be the human resources department, applicant processing employees, or "head hunters" looking for the best people. The information about a subject found online can provide more intelligence than any interview or reference check.

Parents and teachers are encouraged to use this book as a guide to locating social media content posted by children. In many households, the children know more about the internet than the adults. The children use this to their advantage and often hide content online. They know that it will not be located by their parents and teachers, and often post inappropriate content. This book can empower the adults and assist with identifying important personal information.

A large portion of my intended audience is private investigators. They can use this book to find information without possessing a deep understanding of computers or the internet. Explicit descriptions and occasional screen captures will ensure that the techniques can be recreated on any computer. Several universities have adopted this book as required reading, and I am honored to play a small role in some amazing courses related to network security.

I realize that people who use these techniques for devious purposes will read this book as well. Colleagues have expressed their concern about this possibility. My decision to document these techniques came down to two thoughts. First, anyone that really wants to use this information in malicious ways will do so without this book. There is nothing in here that could not be duplicated with some serious searching and time. The second thought is that getting this information out to those that will use it appropriately is worth the risk of a few people using it for the wrong reasons. Please act responsibly with this information.

Custom Search Tool

Throughout this book, I reference several custom search tools that I created to assist with automated queries. I have made available a single repository of every resource discussed in this guide, including the multiple custom search tools. This is presented in an easy to use format with search topics on the left and dedicated query tools within the main area. It can be found at the "Tools" tab of my website **IntelTechniques.com**. This complete archive may be useful as you complete the tutorials within this book. The image below displays the current state of the tool using the custom Facebook search options.



The IntelTechniques Custom Search Tools page.

Finally, a parting thought before you begin your journey through OSINT analysis and collection. This book was written as a reference guide. It does not need to be read straight-through. I encourage you to skip around when needed or if you feel overwhelmed. The second chapter about Linux may make you want to abandon the teachings before ever utilizing an online resource or website. When you encounter material that seems too technical or not applicable, please move on to the next topic. The book is suitable for all skill levels, and there is something here for everyone. You can always return to advanced topics later.

CHAPTER ONE
PREPARE YOUR COMPUTER

The first four editions of this book began with search engine techniques. Right away, I offered my methods for collecting online information from various popular and lesser known search websites. This may have been due to my own impatience and desire to "jump in" and start finding information. This edition will begin much differently. Before you attempt any of the search methods within this book, I believe you should prepare your computing environment.

I was motivated to begin with this topic after teaching a multiple-day OSINT class. On day two, several attendees brought laptop computers in order to attempt the techniques I was teaching during the course. During a break, I observed police officers searching Facebook on patrol vehicle laptops; private investigators using Windows XP while browsing suspects' blogs; and global security professionals looking at hacker websites without possessing any antivirus software or script blockers.

I have also been guilty of all of this. Early in my career of researching OSINT, I did not pay any attention to computer security or proper browsing habits. While I was aware of malicious software, I knew I could re-install Windows if something really bad happened. This was reactive thinking. I believe that we must all proactively attack vulnerabilities in our privacy and security while conducting online research. This chapter is not meant to be a complete guide to computer security or a manual for total privacy. Instead, I hope to quickly and efficiently propose the most beneficial strategies that will protect you from the majority of attacks. Applying the changes mentioned in this chapter will provide a valuable layer of security to your online investigations and overall computing habits. In the next chapter, I present my solutions for guaranteed protection during online investigations.

The most basic place to start is your antivirus. It is likely that most readers already have an antivirus solution and are insulted at the mention of it in a book like this. I will keep my thoughts very brief. If you are using Microsoft Windows, you absolutely need antivirus software. If you are using an Apple computer, you might not. Antivirus applications only protect against known variants of viruses. They do not stop everything. A new virus can often bypass the best software detection solutions. A better defense is applying better browsing habits instead of relying on an application.

There are a dozen popular antivirus companies that will provide a free solution. For most Windows users, I simply recommend to use Microsoft's products. Users of Windows 7 should use Microsoft Security Essentials while Windows 8 and 10 users should use the default Windows Defender included with their installation. Privacy enthusiasts will disagree with this advice, and I understand their stance. Microsoft products tend to collect your computer usage history and analyze the data. Unfortunately, their core operating systems also do this, and it is difficult to

disable long term. Therefore, I believe that Windows users are already disclosing sensitive information to Microsoft. Using their antivirus solutions will not likely enhance the data being collected.

Mac users do not have any built-in antivirus protection, and most do not need any. The software architecture of Mac computers is much more secure, and viruses are rare (but they do still occur). I no longer recommend the free commercial products such as Avast, Kaspersky, and others. They tend to be more of an annoyance than helpful, and their business practices can be questionable. However, I do believe that it is irresponsible to have absolutely no protection whatsoever. When I conduct investigations from a Mac computer, I possess an open-source antivirus solution called ClamAV.

ClamAV (not to be confused with the unnecessary paid option of ClamXAV), is a community-driven antivirus database, which is freely available to anyone. It usually does not score very high on "Top 10 Antivirus" websites, which are usually paid advertisements. However, it is completely free, does not run on your system non-stop, only executes when you desire, and can be completely removed easily. Unfortunately, there is no easy software installation process, and no point-and-click application. You will need to manually update the database through a Terminal command, then scan your system from the same prompt. ClamAV does not remove any viruses, it only discloses the presence and location of suspicious files. In my use, ClamAV has never found a virus that impacted a Mac computer. Instead, it has identified numerous malicious files that target Windows machines, but were present on my system (mostly as email attachments). This notification allowed me to manually remove those files, which could prevent future infection of my Windows virtual machines. If you have concerns about having a "naked" Mac with no antivirus, the following instructions will configure your Mac to be better protected.

First, you must install a package manager called Brew. This program is very beneficial when there is a need to install programs that would usually already be present on a Linux computer. It also happens to have a pre-configured version of ClamAV ready to go. The easiest way to install Brew is to visit the website brew.sh and copy and paste the following command into the Terminal application (Applications > Utilities > Terminal).

```
/usr/bin/ruby -e "$(curl -fsSL https://raw.githubusercontent.com/Homebrew/install/master/install)"
```

After Brew is installed, type the following commands, hitting "Return" after each line, into the same Terminal application used previously.

```
brew install clamav
cd /usr/local/etc/clamav/
cp freshclam.conf.sample freshclam.conf
sed -ie 's/^Example/#Example/g' freshclam.conf
```

These steps will install ClamAV, switch to the installation directory, make a copy of the configuration file, and then modify the configuration file to allow ClamAV to function. You are now ready to update your antivirus database and conduct a scan. Type the following commands into Terminal.

```
ffeshclam -v
clamscan -r -i /
```

The first option will download all virus definition updates, and should be executed before each scan. The second option conducts a scan of the entire computer, and will only prompt you with details of found viruses. While it may appear to be dormant, it is working, and will notify you upon completion. All of these commands must be exact. In order to assist with this, I have created a web page with all of these commands at IntelTechniques.com/clamav. On a final note about ClamAV, you may occasionally receive a false-positive report of a virus. Do not panic. Research the file on the internet and identify the issues. If you receive reports of malicious files within email, simply delete those messages. The use of ClamAV on Mac computers is more about preventing the spread of bad files to Windows users instead of protecting your own machine.

Whether on Windows or Mac computers, protection from malicious software, otherwise known as malware, is vital. Again, there are numerous free options from which to choose. I recommend Malware Bytes for both Windows and Apple users. It is completely free and thorough. I suggest executing, updating, and scanning at least once a week on every device that you use.

- Navigate to http://www.malwarebytes.org/ and select the "Free Download" option.
- Conduct a default installation.
- On a weekly basis, launch the program, update the database, and conduct a full scan.
- Malware Bytes will remove any issues it finds.

Your computer should also be cleaned weekly. As you browse the internet and use applications, unnecessary files accumulate and slow the operating system. I recommend CCleaner for all Windows and Apple users. It is free and easy to use. It provides a simple interface and is used to clean potentially unwanted files and invalid Windows Registry entries from your computer. The following steps will download and install the free version of the application.

- Navigate to http://www.piriform.com/ccleaner/download.
- In the "Free" column, click on "download".
- Execute the program and accept the default installation settings.

After the installation completes, launch the program. You have several options under the Cleaner tab that will allow you to choose the data to eliminate. The default options are safe, but I like to enable additional selections. Clicking on the "Analyze" button will allow the program to identify files to delete without committing to the removal. This will allow you to view the files before

clicking "Run Cleaner" to remove them. If you are running this program on a computer with heavy internet usage, you may be surprised at the amount of unnecessary files present. The first time you use this program, the removal process can take several minutes and possibly an hour. If you run the program weekly, it will finish the process much quicker.

The Registry tab of CCleaner will eliminate unnecessary and missing registry entries. This can help your computer operate more efficiently. The default options on this menu are most appropriate. Click on "Scan for Issues" and allow it to identify any problems. This process should go quickly. When complete, click on "Fix Selected Issues" to complete the process.

The Tools tab provides an easy way to disable specific programs from launching when your computer starts. These programs can slow your computer down when they are running unnecessarily. These can be found by clicking the "Startup" button in the left column. I once selected the Adobe and Java programs and applied the "Disable" button. They were then marked as "No" and would not launch the next time my computer started. If I wanted to reverse this, I could select the entries again and choose "Enable".

Proper antivirus, malware protection, and cleaning solutions will greatly enhance your overall computing experience. It will help your computer to run smoothly and may prevent malicious files from infecting your operating system. It will help protect the integrity of any online investigations. I refer to these steps as the "staples". They are the minimum requirements before proceeding and apply to any computer user.

Those that want to conduct advanced searches on the internet must progress to another level. You must upgrade your web browser and stop relying on Microsoft's Internet Explorer or Edge browsers. I believe that you should only use one of two web browsers: Firefox or Chrome. Many of the techniques in this book, especially in the Application Programming Interfaces (APIs) chapter, will fail when used in conjunction with Microsoft's browsers. They require a more sophisticated solution with proper add-ons. I will focus on Firefox first, as it is my preferred browser for every investigation.

Many readers find that security restrictions on their computers prohibit them from installing any software, including web browsers. While I have found that downloading portable versions of Firefox and Chrome eliminate this restriction, my experience is that this action will upset the computer support personnel that originally enabled the rules. Please research your organization's computer use policies before placing any software on company owned machines.

Those in law enforcement should be more cautious than others. Not only could installing unauthorized software on a government computer violate internal policies, but it could also jeopardize your case in court. If a defense attorney can prove that you violated your own rules and regulations, regardless of how minor or inconsequential, it leaves an opening to request a judge to dismiss your entire findings. Please make sure that you always have the proper authorization to conduct any techniques mentioned in this book.

Firefox (mozilla.org)

The most vital application in this chapter is the Firefox web browser. Most of the search methods that you will learn throughout this book must be conducted within a web browser. Most people settle for Internet Explorer or Edge, which is included with Windows. I do not recommend using those browsers for OSINT analysis. The Firefox browser has enhanced security and a feature called "add-ons" or "extensions". These are small applications that work within the browser that perform a specific function. They will make searching and documentation much easier. I also use, and encourage others to use, the Chrome web browser when necessary. However, many of the extensions that I need are only compatible with Firefox. The following instructions apply to any versions of Firefox, including Windows, Mac, and Linux.

Downloading and installing Firefox is no different than any other application. Detailed directions are readily available on their website. The browser will not look much different from the browser you were previously using. When installing and executing, choose not to import any settings from other browsers. This will keep your browser clean from unwanted data. The next step is to ensure your browser is up-to-date. You can check your version of Firefox by clicking on the Menu button in the upper right (three horizontal lines), then the Help button (?), and finally the option labeled About Firefox. This will open a new window that will display the version of Firefox you are running, or a warning that the version you have is out-of-date.

In November 2017, Firefox released version 57 of their browser. These updates usually go unnoticed by most users, as the changes are minimal. However, this was not the case with 57. This completely new version of Firefox included major speed improvements, a cosmetic face-lift, and most importantly the elimination of legacy extensions (also called add-ons). In the previous edition of this book, I spoke of various Firefox add-ons that would enhance your collection of online information. The majority of these extensions were disabled with this new release. Some developers updated their software to make these options work with the newest version of Firefox while others decided to abandon their projects. At the time of this writing, Firefox is offering an Extended Support Release (ESR) that will safely allow the execution of an older browser which allows legacy extensions. However, this is a temporary solution, and may not be an option by the time that you read this. Therefore, I will only focus on long-term options that should be valid throughout the life cycle of this book.

Before identifying Firefox resources that will aid in our OSINT research, we must first secure our browser to the best of our ability. While the default Firefox installation is much more secure than other browsers, we should still consider some modifications. I personally use Firefox for all of my OSINT investigations, and as my personal web browser. I no longer possess multiple browsers for various tasks. I believe that Firefox is the most robust, secure, and appropriate option for almost any scenario. However, I recommend changing the following settings from the Options (Windows) or Preferences (Apple) menu within Firefox.

General: When Firefox Starts: I choose "Show a blank page" at this prompt. This will make your browser open faster, and eliminate the unnecessary loading of a default web page.

Privacy & Security: Browser Privacy: Deselect the "Remember passwords for sites" and "Use a master password" options. When browsers store a password, they usually do not do so in a secure manner.

Privacy & Security: History: Under the "Firefox will:" option, select "Use custom settings for history" from the pull-down menu. This will allow you to choose everything that is stored or forgotten when you close your browser. Next, uncheck "Remember my browsing and download history" and "Remember search and form history". This will prevent Firefox from remembering any history after your browsing session has closed. Next, check the box that says "Accept cookies from sites". This will allow cookies from the sites you visit. Without cookies, it is very difficult to use social networks, online streaming services, or some search engines. Next, under the "Accept cookies from third party sites" drop-down, select "Never". Under "Keep until", which refers to how long cookies are retained, select "I close Firefox". This option will ensure they are not saved after your browsing session has ended. Finally, check the box that says "Clear history when Firefox closes".

Privacy & Security: Firefox Data Collection and Use: Uncheck both of these options This prevents Firefox from sending data about your session to their servers.

about:config Settings

Firefox allows users to modify many configuration settings, and some of these deal with privacy and security concerns. Though some of these changes can be made in the preferences menu of Firefox's preferences, changes made through about:config tend to be more durable and granular. To access the list of configuration settings, open Firefox and type "about:config" into the URL bar. You will receive a warning about making changes within this area, but the modifications we make will be safe. Choose to accept the risks.

Some of these about:config settings may already be on the "correct" setting, but most probably will not. To change most of these settings you can simply double-click the setting to toggle it between "True" and "False". Some may require additional input, such as a number. Because the list of about:config settings contains hundreds of entries, you will probably wish to search for all of these through the search bar in the about:config interface.

privacy.trackingprotection.enabled: TRUE: This blocks website tracking.

geo.enabled: FALSE: This disables Firefox from sharing your location.

browser.safebrowsing.phishing.enabled: FALSE: This setting disables Google's "Safe Browsing" and phishing protection. If this setting is "true" Google will be able to scan (and store) the sights that you visit for the presence of malware.

browser.safebrowsing.malware.enabled: FALSE: Again, this disables Google's ability to monitor your web traffic for malware, storing the sites you visit.

dom.event.clipboardevents.enabled: FALSE: Many websites will request a notification if you copy text or images from their website. They may also be notified if you select part of a page. This setting disables the ability of websites to access this information. Note that this change may cause issues with copying and pasting text within websites.

media.navigator.enabled: FALSE: Website operators will identify your computer as unique to enable tracking around the web. One such tactic is to track the status of your webcam and microphone (ON/OFF). This disables the ability to website operators to see this information.

dom.battery.enabled: FALSE: Another technique used by website operators to track you is to view your exact battery levels. This setting prevents this information from being shared.

extensions.pocket.enabled: FALSE: This disables the proprietary Pocket service.

WebRTC (Web Real-Time Communications): The next few settings in about:config deal with the WebRTC vulnerability that can allow your IP address to be leaked, even if using a VPN.

media.peerconnection.enabled: FALSE

media.peerconnection.turn.disable: TRUE

media.peerconnection.use_document_iceservers: FALSE

media.peerconnection.video.enabled: FALSE

It is not vital that all of these security settings be applied to your systems. Firefox natively respects your privacy and security more than other browsers. These recommendations are for those that truly want to tweak additional settings that may provide a layer of protection, even if minimal. Next, I will discuss the biggest benefit of Firefox, which is the abundance of helpful browser extensions called add-ons.

Firefox Add-ons (Extensions)

There are thousands of extensions available for Firefox. Some are helpful, some are worthless, and some are just fun. This chapter will discuss thirteen of them. The Firefox add-ons, sometimes called extensions, detailed here will include a website for each option. You can either visit the website and download the add-on or search for it from within Firefox. The former is usually the best way. While Firefox is open, click on the menu in the upper right and then "Add-ons". This will present a page with a search field in the upper right corner. Enter the name of the extension and install from there. The following are my recommendations, in order of importance.

VideoDownloadHelper: Download media from a page with click of a button

Bulk Media Downloader: Download bulk media automatically

FireShot: Generate screenshots of partial and entire web pages

Nimbus: Alternative screen capture for large web pages

uBlock Origin: Block undesired scripts from loading

HTTPS Everywhere: Ensure that you are accessing sites through a secure connection

Exif Viewer: Identify Metadata embedded inside a photograph

MJSONViewer: View API JSON and XML results properly in a browser

User Agent Switcher: Emulate various browsers and devices

Google Translator: Right-click language translation

Image Search Options: Conduct automatic reverse image searches

Resurrect Pages: Enable historical search on deleted websites

Copy All Links: Quickly copy all hyperlinks from a website

The following pages will provide explicit instructions for installing and configuring each of these add-ons. Alternatively, I have configured each of these into a new Firefox browser and exported the settings. If desired, import these configurations into your own Firefox browser for a turn-key solution. This technique will be explained at the end of this section, but I encourage you to consider customizing your own version of Firefox. If you plan to use the Buscador Virtual Machine explained in the next chapter, all of these configurations have already been conducted and are the default option upon boot.

Video Download Helper (downloadhelper.net)

This extension will assist with downloading media that is located during your search. It works well with videos such as those found on YouTube. When this extension is enabled, an icon will appear within your browser that looks like three grey circles. Any time you open a website that includes media content, such as a video, these circles will turn to full color. This is an indication that the media on the page can be extracted. While this add-on will work immediately after installation, I have found specific configuration changes to be helpful to OSINT investigators.

- Click on the icon placed in your menu bar and select the icon for "Settings"
- Click the Behavior tab and change the Max concurrent downloads to 20
- Change the Max Variants to 99
- Select the Hide ADP Variants option

When downloading videos, especially from YouTube, the ADP format requires secondary conversion software to be installed. I do not like this option as it introduces unnecessary software to my machine. Furthermore, I never want to convert video evidence. I simply want to extract the options available directly from the source. Therefore, eliminating the ADP options from our view as explained above reduces the chance of downloading undesired content. In Figure 1.01 (left), the ADP options are present and would not be ideal download choices. In the example on the right, I have eliminated these choices and I am presented with more appropriate options.

You can now extract embedded media files from websites by clicking the icon and selecting the appropriate file. If your desired media is going to be used in court, I recommend downloading all sizes available. If you only want a personal archive, the largest size should be downloaded. You will now have a pure digital extraction of the target video. This is better than a screen capture or recording of the video because there is no loss of data or analog conversion. If downloading a large number of videos, consider the custom script that will be explained in the next chapter.



Figure 1.01: Menu options from Video Download Helper.

Bulk Media Downloader (addons.mozilla.org/firefox/addon/bulk-media-downloader/)

This add-on can make downloading a large amount of media files easy. If you locate a page of several audio or video files, it can be time consuming to save them all manually. Additionally, you run the risk of accidentally skipping a file. Bulk Media Downloader provides a solution. As an example, I navigated to Twitter and searched the word Video. This presented hundreds of embedded videos within a single page. I launched Bulk Media Downloader, which displayed a pop-up option over my browser. In this pop-up, I can select specific file types such as Video or Audio. I chose only the Video option and reloaded the Twitter page in the background. The Bulk Media Downloader tool began populating video links as I scrolled down the Twitter page. Figure 1.02 displays the result. Clicking the Download button retrieved all of the videos in MP4 format as seen in Figure 1.03. This utility works well on sites that have a large number of embedded audio or video files, as well as those that contain numerous documents. You can easily select or deselect entries individually, or select categories at the bottom that fit your needs.



Figure 1.02: A Bulk Media Downloader window.

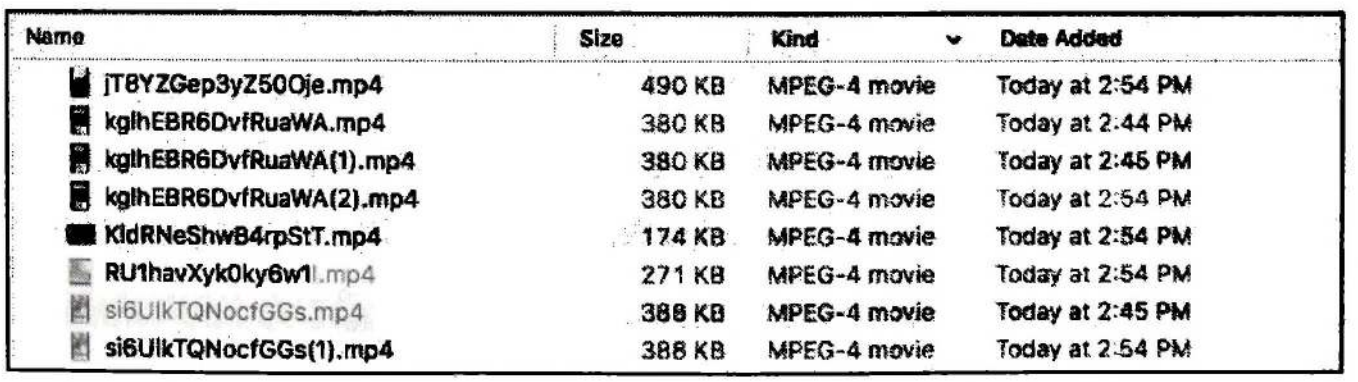


Figure 1.03: Files extracted from Twitter with Bulk Media Downloader.

FireShot (addons.mozilla.org/en-us/firefox/addon/fireshot/)

Documenting and archiving your progress with an OSINT investigation is as important as the intelligence discovered. The general rule is that if you do not have proof of your findings, then they never existed. FireShot provides you with an easy solution to capturing all of your results. When enabled, this extension is a button in the upper right portion of your browser. It appears as a blue square containing the letter "S". Clicking the icon presents a menu with options. The best option is to select "Capture entire page" and then "Save to PDF". This will then create a PDF document of the entire page exactly as it appears in your browser and save it to anywhere you choose. The file can later be archived to a removable storage device. The title of the document will match the title of the web page and it will include the URL of the selected page.

This method is preferred over a standard screen capture for several reasons. A typical screen capture only captures the visible area and not the entire page. You must then open a program into which you "paste" the data and then save the file. The FireShot extension automates this and saves it in a format that is difficult to edit. This can be beneficial during testimony.

By accessing the "Options" area of the menu, you can assign customized naming features. Click "Show filename template settings" in the options page and change the default value to the following.

%n-%u-%t-%y-%m-%d-%H-%M-%S

Be sure to "Apply" and then "Save". This setting will change the default name of each page capture. Each file will be named a numerical value, followed by the website URL, followed by title, and followed by the date and time of capture. Changing the %n value to 0 and the Pad option to 3 will ensure that your captures always start with a numerical value of 0 and ascend chronologically. This can help determine the order of the evidence that you retrieved. Figure 1.04 displays a typical series of results. Notice that you can quickly see the order captured (first three digits), target website, description, and date & time.

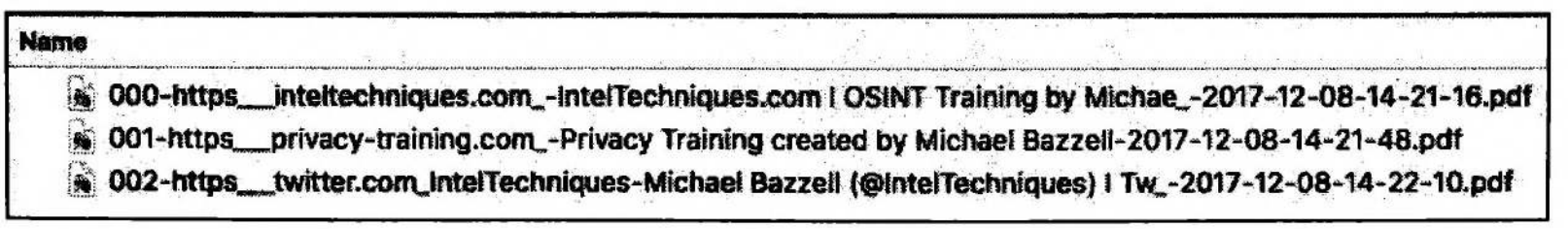


Figure 1.04: Results from FireShot screen captures.

Nimbus (addons.mozilla.org/en-US/firefox/addon/nimbus-screenshot)

While FireShot is my preferred screen capture utility within Firefox, there are some instances where it does not perform well. If you have a target's Facebook page that has a lot of activity present, this may create a screen capture too large for FireShot. The rendering process will likely expend all of the computer's video memory and fail to create the file. When this happens, I use Nimbus as my first backup. Nimbus allows you to specify whether you want to capture only the visible portion of the page, the entire page, or a custom selection from the page. The drop-down menu presents these choices and the result is saved as a PNG file. This is not optimal for online investigations, but is better than no capture at all. Another feature of Nimbus is the ability to manipulate captures. I believe that this is bad practice as we usually want to provide the most authentic and accurate evidence as possible. I do not want to manipulate any potential evidence. Therefore, I recommend the following configurations.

- Click on the Nimbus icon and choose the "gear" icon in the lower-right.
- In the Filename Template, insert {url}-{title}-{date}-{time}. This will name every capture with the URL and title of the target website along with date and time of capture.
- Check the Enable Quick Capture option and select the Entire Page option in the first row and Download option in the second row.

After these changes, clicking the Nimbus icon in the menu bar will no longer present a menu with options. Instead, it will automatically select the entire page, apply the proper file naming, and download the capture as a maximum quality PNG file to your Desktop. While a PDF file created with FireShot is the preferred file format, a PNG file has other advantages. The PNG file is more universal and does not require PDF viewing software such as Acrobat Reader. However, PNG files are easy to edit, and establishing the integrity of the file may be difficult. I believe that Nimbus should be used as a supplement to FireShot.

One common failure of both FireShot and Nimbus is the capture of extremely large Facebook and Twitter pages. While this is rare on computers that have ample resources such as processing power and RAM, it can be quite common on older machines with low specifications. Surprisingly, I have found FireShot to work better on large Twitter profiles and Nimbus to be best for large Facebook pages. I have no logic to offer for this discovery. Again, having both at our disposal will make us better prepared for online evidence collection.

uBlock Origin

In the previous edition of this book, I recommended NoScript as my choice of script blocker. I no longer use NoScript at all. During the transition to Firefox 57, NoScript changed drastically. It became much more convenient to use, at a cost of functionality. We were no longer given granular control of the data that is passed through our browser, and lost several features required for private and secure browsing. I also previously recommended Adblock Plus and Disconnect as privacy add-ons that would help stop unwanted ads, tracking, and analytics. These are no longer present on my systems. I now only use uBlock Origin, as it replaces all three of the previous options. This section may seem a bit overwhelming, but experimenting with the advanced settings should help you understand the functionality. Let's start with the basics.

Install uBlock Origin from the Firefox add-ons page or directly by navigating to the application's website at https://addons.mozilla.org/en-US/firefox/addon/ublock-origin/. You are now protected on a basic level. By default, most known invasive advertisements, tracking code, and malicious content is blocked. This step alone would provide much needed protection from the internet. However, we can take it a step further.

Click on the uBlock Origin icon in the menu and select the Dashboard icon to the right. This will open a new tab with the program's configuration page. On the Settings tab, click the option of "I am an advanced user". This will present an expanded menu from the uBlock Origin icon from now forward. Click on the 3rd-Party Filter tab and consider enabling additional data sets that will protect your computer. I select all options within the Ads, Privacy, Malware Domains, and Annoyances categories. After you have made your selection, click the Update Now button at the top of the page. This will refresh all of the data and apply your new settings. You now have extended protection that will be applied to all visited websites without any interaction from you. When you encounter a web page with a lot of advertisements, such as a news media website, it should load much faster. It will block many of the pop-ups and auto-play media that can be quite annoying when conducting research. This protection will suffice for most users, but dedicated OSINT analysts may choose to take a more advanced approach.

After you have enabled the Advanced settings as explained above, clicking on the uBlock Origin icon should now present an expanded menu that will change as you visit different sites. In order to explain the function of this menu, I will conduct a demonstration using the website cnn.com. Figure 1.05 displays the default view. While this book is printed in black and white, your view will be in color, and likely all options will appear grey. Scrolling down this list of scripts that have either been loaded or blocked, you can see several questionable scripts such as Facebook, Sharethrough, and Turner. These scripts allow tracking across multiple websites and are the technology responsible for monitoring your interests, web history, and shopping habits.

This menu is split into three columns. The first simply identifies the type of code or domain name of the script. The second column is global settings. Anything changed here will apply to all website visits. The third column contains settings for the current website. A single plus sign (+)

indicates that less than ten scripts were allowed from that specific option. Two plus signs indicates that between ten and one hundred scripts were allowed. The single minus sign (-) indicates that between one and nine scripts were blocked from that domain, while the dual minus signs tell us that ten to one hundred scripts were blocked. In Figure 1.05, we know that over ten scripts were allowed to run from cdn.cnn.com, and at least one script was blocked by smetrics.cnn.com. This is all default behavior and provides a balance of functionality and security. uBlock Origin decides which content should be allowed and which should be blocked.

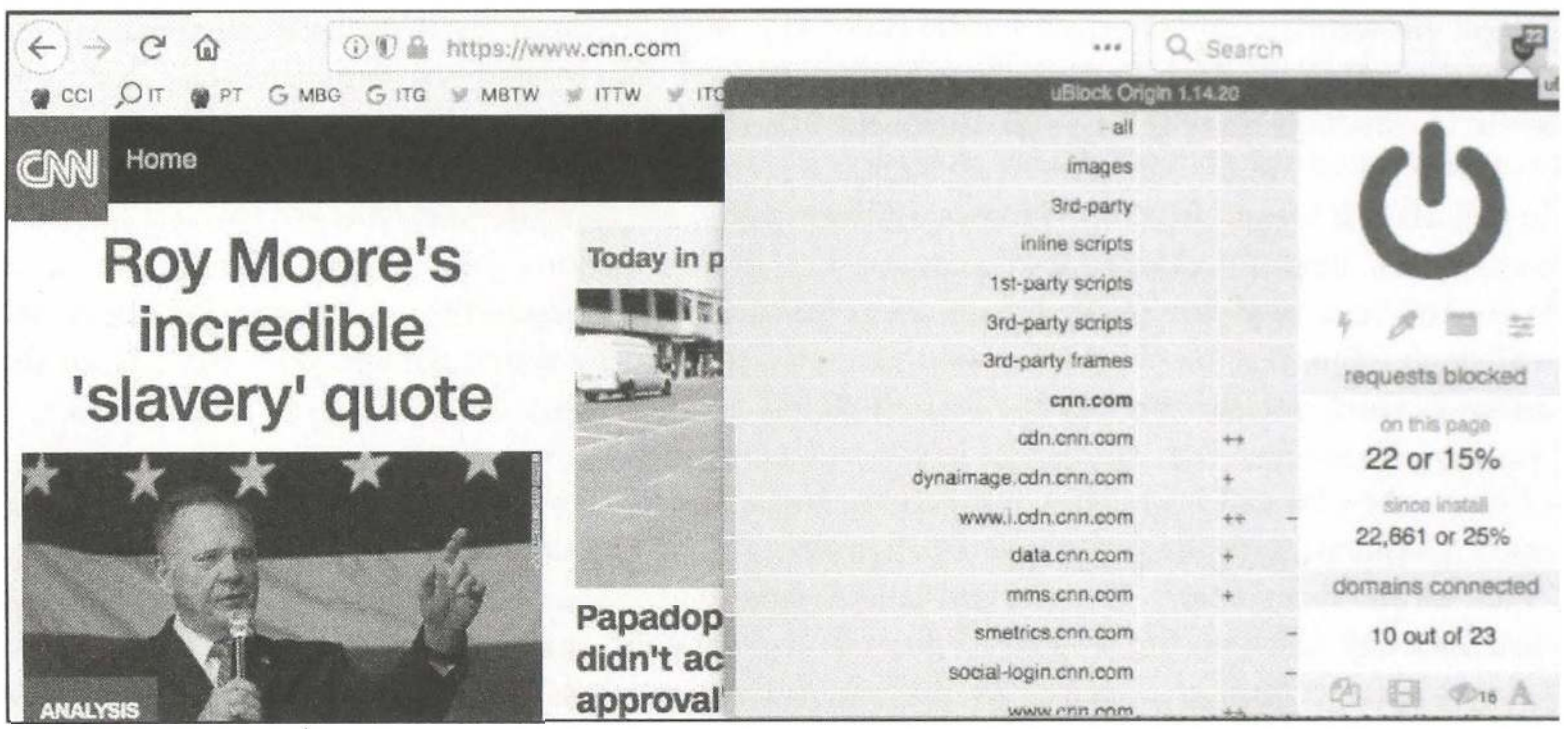


Figure 1.05: An advanced view of uBlock Origin.

Using this same page, let's modify the options. In Figure 1.06 (left), I have clicked on the far-right portion of the first cell in the third column. This turned the entire third column red in color. This action activated an option to refresh the page (middle arrows) and an option to save the change (upper left "padlock"). Clicking the padlock and then refreshing the page presented me with the example in Figure 1.06 (right). Since I blocked every script, the page would not fully execute. It could not load images, design scripts, or any JavaScript. This is not useful at all, so I disabled my actions by clicking on the middle section of the top cell in the third column, which turned the entire column back to grey in color. Saving these changes and refreshing the page brought me back to the example in Figure 1.05.

We can also take this to the opposite extreme. In Figure 1.07 (left), I clicked on the far-left portion of the top cell in the third column. This turned the entire column green in color, and allowed all scripts to load on cnn.com. This includes the dozens of intrusive scripts that could load advertisements on the page. You can also see that small plus signs confirm that scripts were allowed to run while the minus signs in Figure 1.06 (right) state the opposite. For most users, this allowance would seem irresponsible. However, there is a specific reason that we want the ability to allow all scripts. If you are collecting evidence, especially in criminal cases, you may

want to archive a page exactly as it was meant to be seen. When we block scripts, we are technically modifying the page (evidence). By intentionally allowing all scripts before the collection of the screen capture, we know that we are viewing the page in an unmodified format. This may be overkill for many investigators, but you should know your options.

Next, we will modify the second (middle) column, which will apply the settings globally. By default, all options are grey in color. This indicates that the default block list is applicable, and only invasive scripts will be blocked everywhere. I clicked on the far-right portion of the top cell in the second column. This turned the entire column red, and indicates that all scripts across all websites will be blocked. After I saved my changes, every website will only load the most basic text content. This will prohibit much of our research.

Loading a page such as a Twitter profile resulted in no useable content. By clicking on the uBlock Origin icon and clicking the middle sections of specific cells within the third column, I enabled those scripts without allowing everything on the page. While you cannot see the colors in Figure 1.07 (right), you can see the difference in shading. In this example, the entire second column is red. This indicates that all scripts are blocked globally. The third column is mostly red, but the options for twitter.com, twimg.com, and others are grey. Those scripts will be allowed, if approved by uBlock Origin's rules, only for that domain. If I load a blog that has scripts from Twitter, they would still be ignored.

These are extreme examples. Let's bring this back to some sanity. The following is how I recommend using uBlock Origin. Install, enable advanced options, and proceed with your work. When you arrive at a website that is blocking something you want to see, open the menu, and click on the far-left section of the top cell in the third column. That will allow everything to load on that page, and that page only. When you are about to navigate to a questionable site that may try to install malicious code on your machine, click on the far-right section of the top cell in the second column. That will block all scripts on all pages. Conduct your research and reverse the change when you are finished. Remember to click the save button (padlock) after each change.

Hopefully, you are practicing these settings and learning how this program functions. It is an amazing option that has protected me many times. If you are doing things right, you have likely completely messed-up your settings and are now blocking things you want while allowing things you do not. Don't worry, we can reverse all of our mistakes by first making the global (second column) settings back to grey (middle section of top cell). Next, return to the dashboard settings of the add-on, and click on the My Rules tab. In the second column (Temporary Rules), click Edit, highlight all of your customizations, and delete them. Click the Save button in this same column and then the Commit button to apply these settings everywhere.

The huge benefit of uBlock Origin over other options is the simple ability to block malicious scripts without customization, while having an option to allow or block any or all scripts at our disposal. This is a rarity in these types of add-ons.

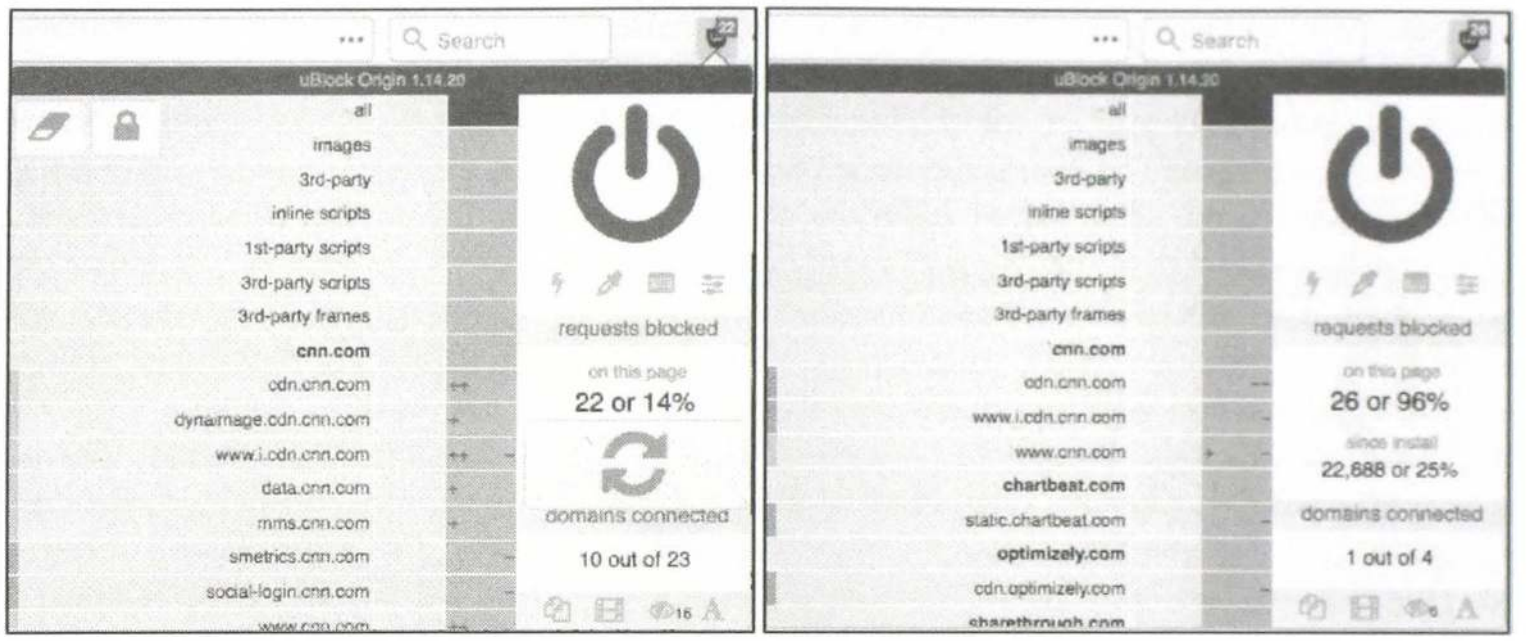


Figure 1.06: Disabled scripts within uBlock Origin.

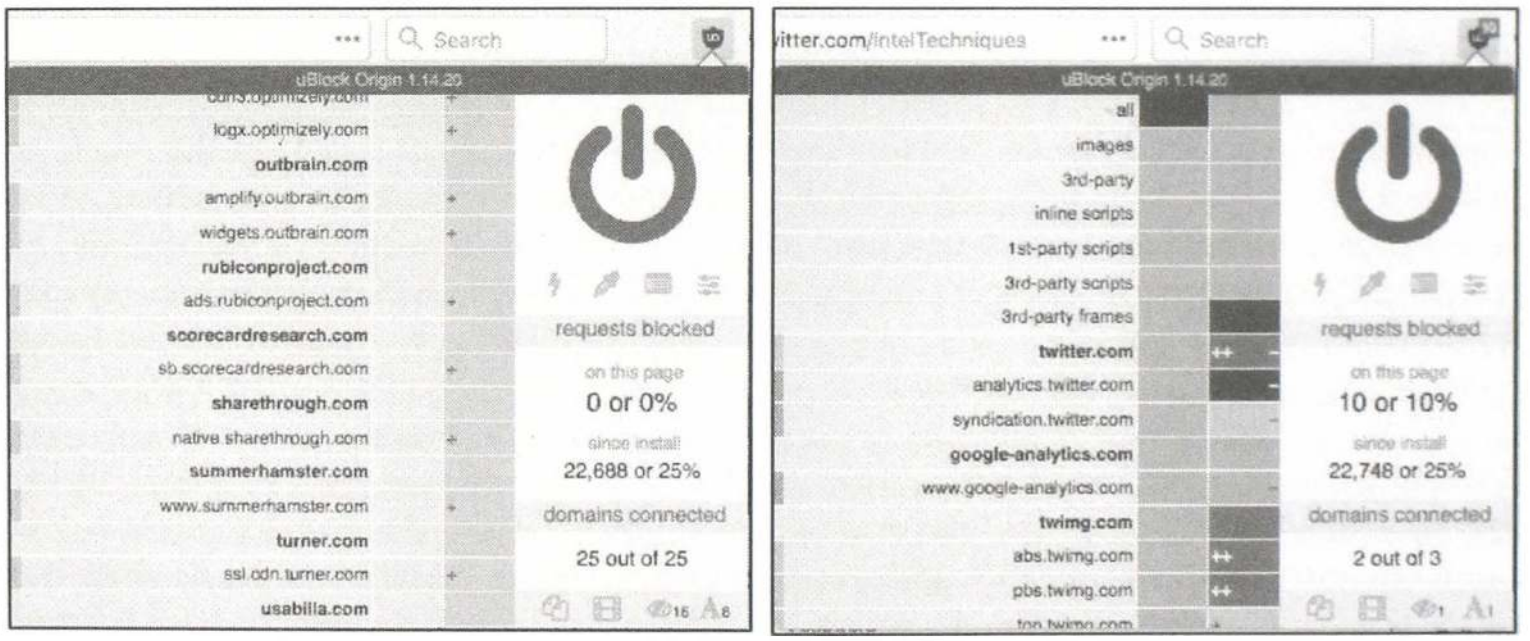


Figure 1.07: Fully and partially enabled scripts with uBlock Origin.

HTTPS Everywhere (addons.mozilla.org/en-us/firefox/addon/https-everywhere)

This extension encrypts your communications with many major websites, making your browsing more secure. It is produced as a collaboration between The Tor Project and the Electronic Frontier Foundation. Many sites on the web offer some limited support for encryption over HTTPS, but make it difficult to use. As examples, a site may default to unencrypted HTTP, or fill encrypted pages with links that go back to the unencrypted site. The HTTPS Everywhere extension fixes these problems by rewriting requests to these sites to HTTPS. This happens automatically after installing this add-on, and you do not need to take any additional actions. If the icon for this extension crowds your browser, you can remove it by clicking the Firefox menu, then Customize, then dragging the icon away from the menu bar. It does not need to be visually present to function.

Exif Viewer (addons.mozilla.org/en-us/firefox/addon/exif-viewer)

This extension provides right-click access to the Exif data embedded into images. Chapter Fourteen explains what Exif data is and how it can be useful. With this extension enabled, you can right-click on any full size image located on a web page. The menu option is "View Image Exif Data" and a new window will open when selected. This window will identify any available metadata about the image. Figure 1.08 (left) displays the right-click menu with a new option to View Image Exif Data. Figure 1.08 (right) displays partial results that identify the make and model of the camera used to capture an image.

Overall, most photos on social networks do not contain any metadata. They have been "scrubbed" in order to protect the privacy of users. However, many blogs and personal websites display images that still contain metadata. While Chapter Fourteen will explain online websites that display this data, a browser add-on is much more efficient. In my experience, this extension will increase the amount of times that you will search for this hidden content.

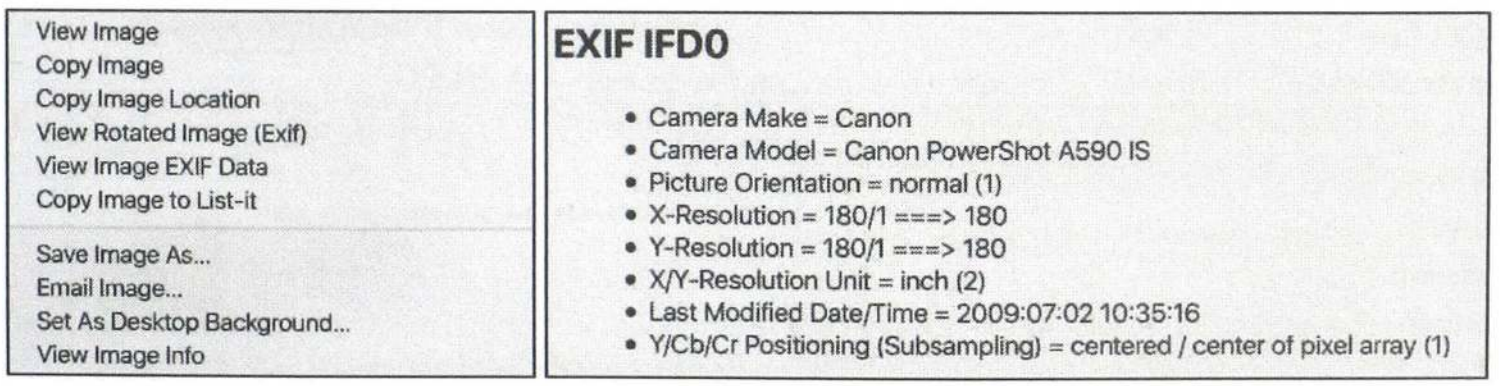


Figure 1.08: The right-click menu and result from an Exif Viewer search.

MJSONViewer (addons.mozilla.org/en-US/firefox/addon/mjsonviewer/)

This extension will probably go unnoticed, which is a good thing. MJSONViewer allows JSON and XML files to be opened and viewed within a web page instead of saving them and viewing the files in a text editor. These files are returned when querying specific types of data such as server content. This mostly applies to the Application Programming Interfaces (APIs) discussions in Chapter Twenty. Many of these APIs deliver the results in a view that is not intended for web browsers, even Firefox. Your browser may try to download a file instead of displaying the content. With this extension installed, the API results should appear within the browser every time. Without this extension, some of the API searches will not function. There is nothing to configure with this add-on. If your results during Chapter Twenty and others appear scrambled or completely missing, install this add-on. Below is a partial output from my Facebook video tool that appears crammed together.

{"videoID":"1498736236888792","lat":-25.457977526294,"long":-49.21751453494,"name":"-Can\u00e7\u00e3oNova Curitiba","startTime":1512773369,"previewImage":"htt-ps:\/\/-scontent-lax3-1.xx.fbcdn.net\/v\/t15.0-10\/s480x480\/24519329_1498758-5535532-27_-3265374544642179072_n.jpg?oh=db1c9566bf696f1b57560ecd78beae4b&oe=5ACDB656","viewerCount":72,"formattedCount":"72","publisherCategory":"Religious Organization",-"profile-Picture":"https:\/\/scontent-lax3-1.xx.fbcdn.net\/v\/t1.0-1\/p200x200\/-16003168-_1181-861118576307_6711915926513163297_n.jpg?oh=9220f094e016b60d6a99db7a30fabb3d&oe=5ACDEF67","width":480,"height":360,

Below is the same result after installing the MJSONViewer add-on.

"videoID": "1498736236888792",
"lat": -25.510737280581,
"long": -49.267747889626,
"name": "Canção Nova - Curitiba",
"startTime": 1512773369,
"previewImage": "https:\/\/scontent-lax3-1.xx.fbcdn.net\/v\/t15.0-10\/s480x480\/24243481_1498757410220008_7440998993137500160_n.jpg?oh=d1c8db78d3338060472c6087106f8316&oe=5AD2089F",
"viewerCount": 71,
"formattedCount": "71",
"publisherCategory": "Religious Organization",
"profilePicture": "https:\/\/scontent-lax3-1.xx.fbcdn.net\/v\/t1.0-1\/p200x200\/16003168_1181861118576307_6711915926513163297_n.jpg?oh=9220f094e016b60d6a99db7a30fabb3d&oe=5ACDEF67",
"width": 480,
"height": 360,

User Agent Switcher (addons.mozilla.org/en-US/firefox/addon/user-agent-switcher-revived)

Occasionally, you may visit a website that does not want to cooperate with Firefox. Browsers notify websites of their identity and websites can alter or refuse content to certain products. One example is that some older websites require Microsoft's Internet Explorer to view the content. Even though Firefox is capable of displaying the information, a website can refuse the data to the browser. Another example is mobile websites that display different content if viewing from an iPhone instead of a computer. This can now all be controlled with User Agent Switcher.

When installed, you have a new option in your browser. The menu allows you to choose a mobile operating system, such as iOS or Android, or a desktop browser such as Internet Explorer or Chrome. It will also allow you to specify your operating system such as Mac or Windows. Whatever you choose, this data will be sent and confirmed to any site that you visit. If you visit a website of a tech-savvy target, he or she may know that you were looking around. You may also be revealing that you are using a specific browser, such as Firefox, and a Windows computer (common in government). You could now change your agent to that of a mobile device or Google Chromebook which may not look as suspicious. To do so, you would click on the menu bar icon, and simply select the desired configuration. To return to the default Firefox option in your native operating system, click on the checkmark icon in the lower left. Figure 1.09 displays an example where a mobile version of Yahoo was delivered to a desktop computer.

I have used this on several occasions to bypass poor security protocols. During one investigation, I had encountered a web forum of a hacking group that always appeared blank on visit. Google had indexed it, but I could not see any content. By changing my default agent to Firefox on a Linux machine, I was allowed to see the content. The group had enabled a script that would only allow the page to be viewed on Linux computers. While still employed by the government, various mandated online training needed to be completed in order to maintain specific certifications. This government-hosted training was poorly designed and required users to access via Internet Explorer. Since I used an Apple computer, I could not connect until I changed my agent to Internet Explorer within my Firefox browser.

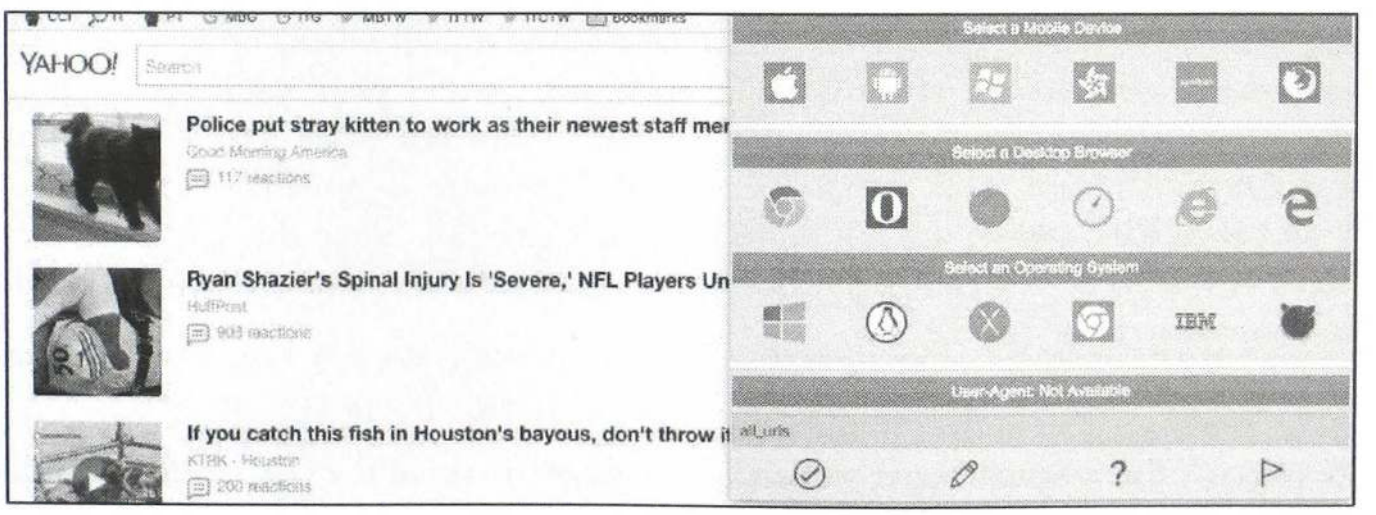


Figure 1.09: User Agent Switcher disguising a desktop system as a mobile device.

Google Translator (addons.mozilla.org/en-US/firefox/addon/google-translator-for-firefox/)

If you find yourself using Google's free language translation service to translate foreign text into English, this extension will be a welcome addition to your Firefox browser. When installed, any text you highlight and right-click will present a new menu option. This will automatically translate the text into English without leaving the target website. This will work on any site including social networks. Figure 1.10 displays a Twitter post translated to English through the extension. The right-click menu will display "Translate selection with Google Translate".



Figure 1.10: The Google Translator Extension in use.

The importance of this extension is not that it can translate text. Anyone can copy and paste text into Google Translate to understand the content. The benefit here is that it delivers the translation back into the original website. Lately, when using Google Translate for an entire page, the final formatting is much different than the original. I have found myself relying on this add-on in order to preserve the original layout of the content, such as a Twitter post, while also displaying the translation of the text. Additionally, Google Translate does not work well when submitting a URL of a Facebook page. This is because Google Translate is not logged into a Facebook account, and therefore cannot see the same view that you can. Since this add-on can translate text within a visible page, you can easily convert foreign Facebook posts into English translations based on your view of the profile.

Clicking the Translate icon in the Firefox menu will open a new tab and present an official Google Translate page of the original target website. This shortcut alone may be worth the installation for those that often deal with foreign language websites. There are numerous Firefox translation add-ons, and you should consider finding the best option for your investigations.

Image Search Options (addons.mozilla.org/en-US/firefox/addon/image-search-options/)

A later chapter explains reverse image search engines and how they can identify more target photographs. Popular options include Google Images and Tin Eye. This extension automates the reverse search when an image is right-clicked. When installed, "Image Search Options" is present when you right-click on an image. Highlighting this option presents several reverse image search services including Google, Bing, TinEye, Yandex, Baidu, and others. You will later learn how my online search tool will execute an image search across all of those services at once. However, this tool can be beneficial due to the convenience and obscure services such as Karma Decay, which looks for copies of images on Reddit. This add-on removes any excuse to not always check reverse images on target websites. With this add-on enabled, you will be ready to enhance your searches during that investigation. Figure 1.11 displays the options after right-clicking an image located on a target website.

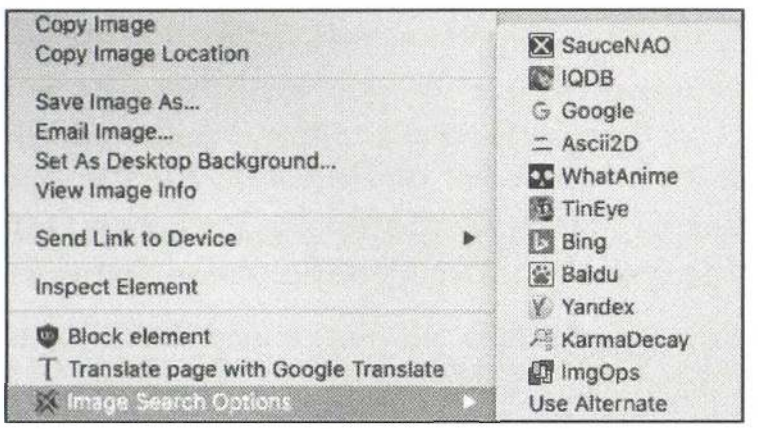


Figure 1.11: A reverse image search menu.

Resurrect Pages (addons.mozilla.org/en-US/firefox/addon/resurrect-pages/)

This extension provides a link to archived versions of websites whenever a page has been modified, is unavailable, or has been deleted. Right-clicking on any site will offer a new menu option of "Resurrect this page". That option will present the following archive services.

Google Cache: A standard cache of the target address from Google
Google Cache Text: The text-only view of a standard Google cache
The Wayback Machine: A link to the target page within The Internet Archive
Archive.is: Any captures of the target address domain on Archive.is
WebCite: Any captures of the target address domain on WebCite

This add-on will not give you any content that you could not locate manually from these sources. Instead, it serves as an easy way to quickly identify interesting content. You will learn more about online archives later in this book.

Copy All Links (addons.mozilla.org/en-us/firefox/addon/copy-all-links/)

This simple add-on will identify any hyperlinks within the source code of an individual web page. It will store the links within your operating system's clipboard, which will allow you to paste them into any application of your choice. While only a small utility, it can quickly turn a large project into an easily completed task. After installing, I recommend navigating to the "options" menu within the add-on and selecting "Copy only clickable links". This will eliminate many unwanted results.

Using the utility is fairly straightforward. While on any website, right-click anywhere in the page and highlight the "Copy All Links" option in the menu. This will present you with options to either copy the links in either the current tab or within all open tabs of your browser. You can then choose to copy all file sharing links, direct file links, or all links. I always recommend the last option. The links will be stored in your clipboard and you can paste them into Notepad, Excel, or any other productivity application. There are unlimited uses for Copy All Links, and below are a few of my favorite.

Facebook: When I am on my target's list of Facebook friends, I will use Copy All Links to quickly record each hyperlink to an individual's profile. I will then paste these into Excel for later analysis. Comparison with previous captures identifies those that were "unfriended".

Twitter: When I am viewing a Twitter profile, I will use this utility to capture all links to external websites and photos.

YouTube: When viewing a person's YouTube videos page, Copy All Links allows me to paste the entire link to every video into a report.

eBay: While viewing results from a search for a specific fraudulent product, I can quickly copy the active hyperlinks to each auction and paste them directly into a report in seconds.

Backpage: While viewing ad results for suspected human trafficking victims, I can copy all active hyperlinks and paste directly into report, email, or memo for other investigators.

Documents: When I encounter a public FTP server or open web directory (Chapter Three), this tool allows me copy the native links to all files encountered. This is helpful for documentation after downloading all of the data.

Performing screenshots during my investigation in these examples would never identify the direct links to the visible content. Only hovering over the links would temporarily identify the source. A combination of screen captures and link collection with this add-on provides a much more comprehensive report.

Importing a Pre-Configured Profile

At this point, it may seem overwhelming when thinking about the abundance of add-ons and their proper configurations. I currently use several Windows, Apple, and Linux virtual machines and must keep my Firefox browser updated on all of them. I no longer manually update each browser. Instead, I maintain a single Firefox browser that includes all customizations that I desire. I then import these settings into any other browsers in order to replicate the experience across every computer that I use. The following instructions allow you to use this same customization file for import into your own investigative computer. The details will vary slightly based on your operating system and version. Only execute this tutorial on a new install of Firefox that has no saved settings, or within a version of Firefox that you want to be overwritten. Do not overwrite your current version if you have bookmarks, extensions, or other data that you want to keep. You should backup any settings if proceeding on an older install. As a final warning, the following steps will overwrite any custom options applied to your Firefox installation.

- Navigate to the following website and download the compressed zip file. Unzip the file, which contains a folder called Profile, to your Desktop.

  https://inteltechniques.com/data/profile2018.zip

- Open your version of Firefox and click the menu button (three horizontal lines), click the help question mark, and then select Troubleshooting Information. The Troubleshooting Information tab will open.

- Under the Application Basics section, click on Open (or Show) Folder. A window with your profile files will open. Close Firefox.

- Paste the content of the downloaded Profile folder into this folder. Overwrite any files when prompted. Restart Firefox.

The result should be a copy of Firefox that contains every add-on mentioned in this chapter. The configurations discussed here have already been applied and you are ready to begin your searching. This profile could be copied to an unlimited number of computers. While I will attempt to update this file occasionally, it will likely be somewhat outdated when you install. Correct this by clicking the settings button on the Firefox add-ons page and choosing Check for updates. This will apply any add-on updates available. You may also notice a row of bookmarks saved within this profile on the Bookmarks bar. These are actually JavaScript programs that perform a function when executed on a target page. The following explains each option as of this writing.

JavaScript Bookmarklets

A bookmarklet is a bookmark stored in a web browser that contains JavaScript commands that add new features to the browser. Bookmarklets are unobtrusive JavaScripts stored as the URL of a bookmark in a web browser or as a hyperlink on a web page. Regardless of whether bookmarklet utilities are stored as bookmarks or hyperlinks, they add one-click functions to a browser or web page. When clicked, a bookmarklet performs one of a wide variety of operations, such as running a search query. For example, clicking on a bookmarklet after selecting text on a webpage could run an internet search on the selected text and display a search engine results page. The following are included with the previously mentioned Firefox profile available for free download. They each execute specific code against the target website within the browser. If you do not see these new bookmarklets after applying the downloaded profile mentioned previously, click on View > Toolbars > Bookmarks Toolbar within Firefox.

Modified: Displays the date that a static website was last modified

Cached: Displays a Google Cache of the website (Chapter Three)

Wayback: Displays the website archive within the Wayback Machine (Chapter Three)

Source: Displays the source code of a website

Site: Conducts a Site: search of the website (Chapter Three)

YTReverse: Conducts a reverse image search of a video (Chapter Fifteen)

Notes: Opens a blank page for note taking

FB-ID: Displays Facebook User ID of current profile (Chapter Four)

FBExpand: Expands all comments on a person's Facebook Timeline (Chapter Four)

ClearTD: Removes any Tweets populated within TweetDeck (Chapter Five)

TorView: Opens a Tor website through a proxy for view natively (Chapter Three)

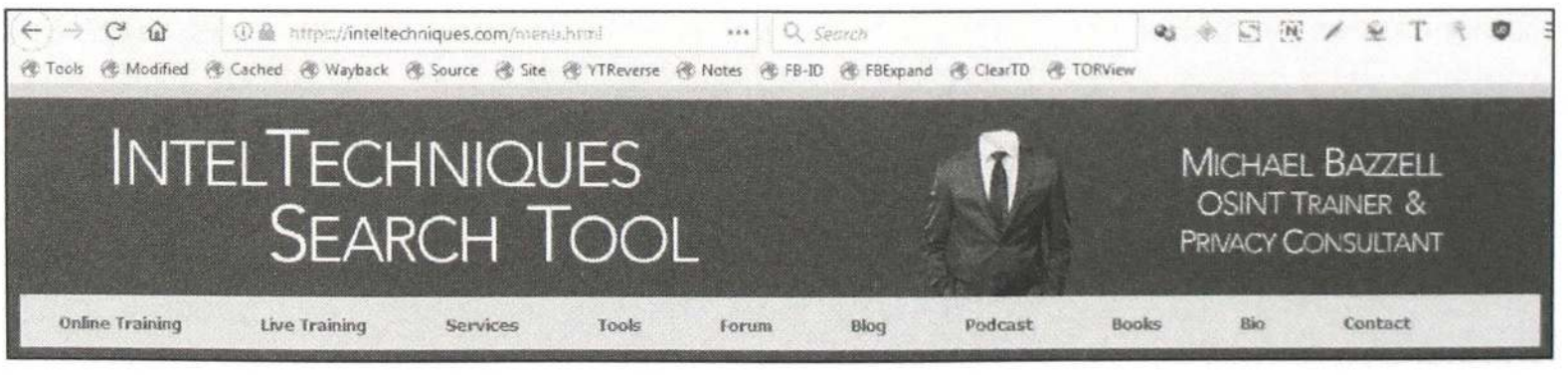


Figure 1.12: A row of JavaScript Bookmarklets available for download.

Chrome (google.com/chrome/browser/desktop)

Chrome is an excellent browser that is known for being very fast and responsive. Chrome is also very secure by nature, but compromises privacy since Google receives a lot of data about your internet usage. Both Firefox and Chrome "sandbox" each tab. Sandboxing restricts the content in that tab to that tab only, preventing it from "touching" other tabs in the browser, or the computer's hardware. This is a very important feature in preventing malware from being installed when you visit a malicious website.

While I always prefer Firefox as my browser for investigations and daily usage, Chrome is my browser for training events. This is due to stability when loading dozens of tabs, and your system should have a lot of RAM if you want to take advantage of Chrome's power. For investigative purposes, Chrome can use several of the add-ons previously mentioned for Firefox. I highly recommend uBlock Origin as discussed previously on any browser that you use, including Firefox, Chrome, Safari, and Opera.

The only time that I use Chrome during an investigation is when I am forced because of a Chrome-specific utility. There are a few extensions for Chrome that will not work on Firefox. I will focus on those here. Before discussing any investigative resources, I suggest you harden your Chrome security. Enter the Settings menu and consider the following changes.

Privacy: Beside the content settings button is a button labeled "Clear browsing data…" This button will open a dialogue that allows you to clear any or all of data stored during your sessions. You may erase information for a period of time ranging from the last hour to "the beginning of time". You may wish to use this function to clear all of your browsing data daily. Alternatively, you could install the extension Click&Clean to automate this chore.

Passwords and forms: I recommend disabling these features by unchecking both boxes: "Enable Autofill to fill out web forms in a single click", and "Offer to save your web passwords". If you have stored form-fill information or passwords in Chrome, I recommend removing any data before conducting investigations.

Chrome Extensions: To install add-ons in Chrome, navigate to the settings menu. Click "Extensions" on the upper left side of the Chrome interface. You will be presented with all the add-ons that are currently installed in Chrome. I recommend uninstalling any add-ons that you did not personally install or research for trustworthiness. Furthermore, most extensions previously explained for Firefox can be installed in the same manner in Chrome. The following Chrome-only extensions may provide additional benefit to your online research.

Prophet (recruitingtools.com/prophet)

Prophet monitors the social networks that you visit and supplies additional details about the targets that you are researching. It does not require an account. After installation, launch Prophet by clicking on the black arrow button in the upper right of your browser. This works best when you are actively on the social network profile of your target. Figure 1.13 displays the view while launched from a person's Twitter page. The results identify her AboutMe, Facebook, FourSquare, Google+, LinkedIn, and Klout profiles. It also connects directly to her personal blog and Flickr page. The "Find Email Address" option reveals two verified email addresses that belong to the target.

I have successfully used this extension on numerous investigations. While the methods that Prophet uses to obtain the data can be replicated with manual searching, it is a laborious process. This extension saves time. In one investigation, I needed to quickly locate the Facebook pages connected to several Twitter profiles involved in a threat case. Clicking through each profile, with the Prophet sidebar expanded, immediately identified the majority of the accounts. A two-hour task was completed in less than fifteen minutes. This tool works best when executed from a Twitter, Facebook, Google+, or LinkedIn profile. It does not work well from blogs or personal websites.



Figure 1.13: A Prophet search result from a Twitter profile.

360Social (https://www.360social.me/)

Similar to Prophet, 360Social aims to immediately discover social accounts associated with the current target page. Once installed, this extension resides in the Chrome menu. When you visit any social network profile, the icon will switch from greyscale to color, indicating potential information. Clicking the icon presents the full menu as seen in the right-side portion of Figure 1.14. In this example, the same target was chosen as was used previously. However, much more information was obtained. We now have direct links to her Twitter, Google+, AboutMe, Flickr, Foursquare, Github, Instagram, Klout, Yelp, and YouTube profiles. We also have links to her personal blog, employer, Snapchat account, and IMDB page. As a further benefit, hovering over any social network profile link, such as her friends or retweets, will immediately change the detail menu to information found about the new target associated with the link.

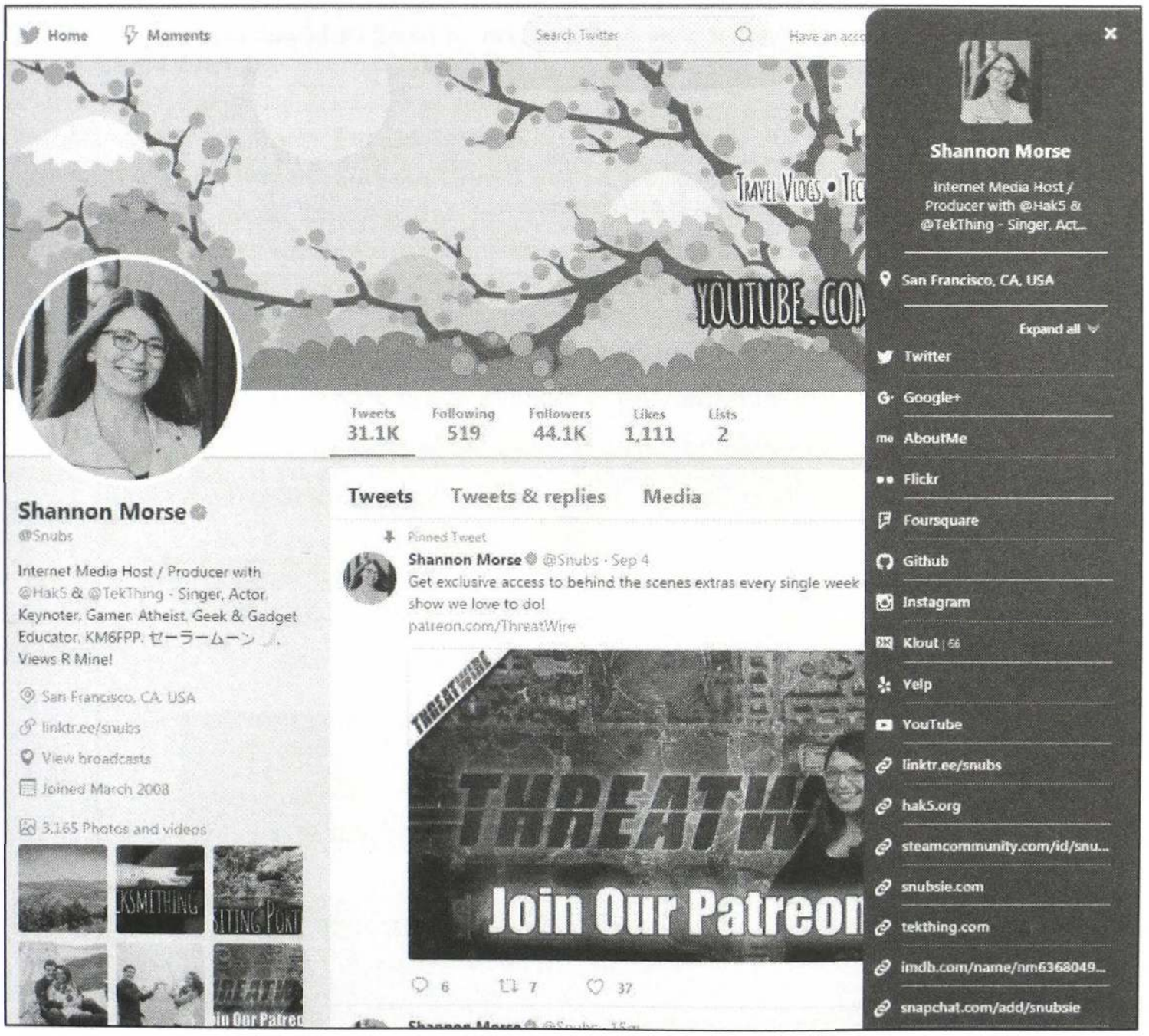


Figure 1.14: A 360Social detail view from a target Twitter account.

Hunchly (hunch.ly)

While FireShot and Nimbus were explained as free options that work with both Firefox and Chrome, they both have their limitations. Neither work extremely well with large social network profiles and both provide no type of file management solution. While I try to focus only on free resources, this book would not be complete without a discussion about Hunchly. Hunchly is a paid tool that is designed to optimize your data capture and analysis during an OSINT investigation. Hunchly takes full content captures of every page that you visit so that you don't lose information during the course of your investigation. Additionally, it automatically does the following:

- Creates a cryptographic signature for each page captured for verification purposes
- Automatically extracts EXIF metadata from every photo encountered
- Enables you to tag pages for easy organization of small or large cases
- Powerful full text search of all captured pages and EXIF data
- Flexible export and reporting options
- Automatic attachment of downloads including documents and video files
- API integration with tools such as Maltego

Hunchly is completely integrated with Google Chrome so you can stay in your browser while you are doing your investigative work. With Hunchly working in the background, you never have to worry about remembering to take screenshots or annotate with some tool. All of the pages are captured, timestamped and documented automatically.

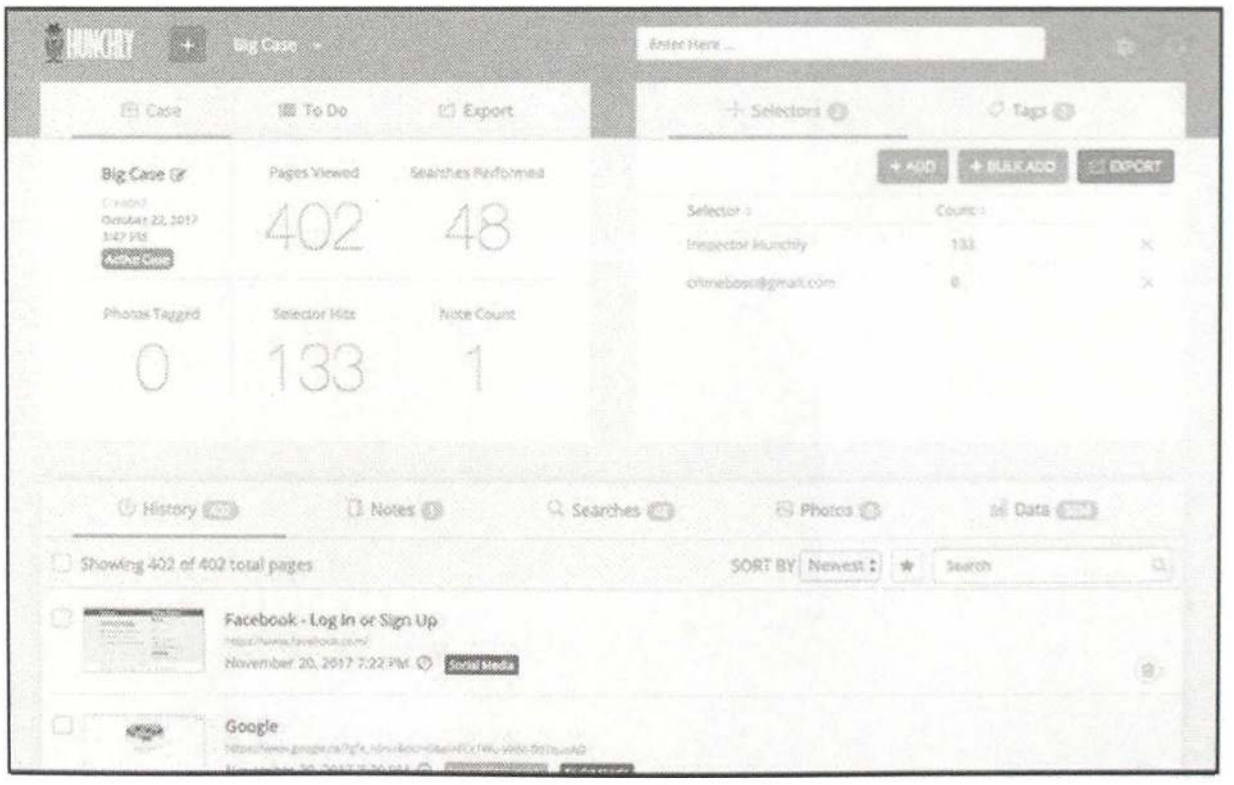


Figure 1.15: The Hunchly capture tool in use.

Tor Browser (torproject.org)

Tor is an acronym for The Onion Router. Basically, it allows you to mask your IP address and appear to be browsing the internet from a false location. Normally, when you connect to the internet and browse to a website, that website can identify the IP address that was assigned to you from your internet service provider. This can often identify the city and state that you are in and possibly the business organization where you are currently located. In some instances, it can identify the building you are in if you are using public wireless internet access. The owner of the website can then analyze this information which may jeopardize your investigation. This is one of the many reasons that I recommend the uBlock Origin add-on for Firefox which was explained earlier. uBlock Origin will block most of the analytic code within websites that monitors your information, but it will not stop everything. Occasionally, you may want to change your IP address to make you appear to be someone else in a different country. This is where Tor excels.

The Tor bundle available for free download is completely portable and requires no installation. After download, unzip the file and extract all of the data. You are now ready to start the program by double clicking the "Start Tor Browser" icon. The first task that Tor will complete is to create a connection to a Tor server. This connects you to a server, usually in another country, and routes all of your internet traffic through that server. After the connection is successful, it will load a custom version of the Firefox browser. Now, every website that you visit through this browser will assume that you are connecting through this new IP address instead of your own. This provides a layer of privacy to stay hidden from a suspect. This may be overkill for most investigations. If you are only searching and monitoring common services such as Facebook, Twitter, or YouTube, this service is not needed. If you are visiting personal websites and blogs of a tech savvy hacker, you should consider Tor. When using Tor, you may notice a drastic decrease in the speed of your internet. This is normal and unavoidable. This often improves the longer you are connected. To stop the service, simply close the browser. This will disconnect the Tor network and stop all services. Figure 1.16 displays the IP address assigned to me through the Tor Browser (top) and a browser not using Tor (bottom). Any activity conducted through the Tor browser is not associated with my real internet connection an appears to be originating in Canada.

IP Address	216.239.90.19 [Hide this IP with VPN]
IP Location	Montreal, Quebec (CA) [Details]
Proxy	216.239.90.19, 198.143.60.25

IP Address	209.58.129.99 [Hide this IP with VPN]
IP Location	San Jose, California (US) [Details]
Proxy	209.58.129.99, 198.143.34.33

Figure 1.16: A Tor IP address and location (top) and actual data (bottom).

Virtual Private Network (VPN)

I have one last topic to present before proceeding to actual search techniques. It is a technology that I strongly support, and carries a minimal cost. I believe that every OSINT researcher should possess and use a virtual private network (VPN) at all times. A VPN extends a private network across a public network, such as the internet. It enables users to send and receive data across shared or public networks as if their computing devices were directly connected to the private network, thus benefiting from the functionality, security and management policies of the private network. Similar to the Tor browser, which was mentioned earlier, a VPN masks your identity online. Two specific examples should help demonstrate the need for this resource.

If you are on your home computer, and connected to the internet, you are using a connection from your internet service provider (ISP). If you navigate to a website that is monitoring visitors, it knows your IP address, approximate location, and internet provider and type (cable, DSL, etc.).

However, if you are on that same computer, with the same internet connection, you can use a VPN to protect you. The VPN software connects your computer to one of their servers over the internet connection. This encrypted traffic cannot be deciphered by the ISP. When your traffic gets to the VPN server, it sends and receives your data, returning incoming packets to you. The websites that you visit believe that you have the IP address of the VPN server. They do not know what type of internet connection you possess nor your location.

Some readers may wonder why they cannot simply use the free Tor service for this scenario. While you could, it is not always advised. Tor connections can be too slow for constant use. Also, some websites will not let you access their services through a Tor proxy. Connecting to your bank through Tor will likely set off alarms, and may prevent you from access. I believe that Tor is great when you truly need to hide your entire connection. I believe that every day browsing is better suited for a VPN.

If you work for a large corporation, you may already have access to a corporate VPN. Ask around and identify your options. Those that do not have a corporate solution will need to purchase a VPN service. While there are a few providers that give away free VPNs, I never recommend them. They are extremely slow and often use your internet connection for other people's traffic. Instead, consider purchasing access from my favorite provider, PIA.

Private Internet Access (PIA) is a commercial VPN that is very popular. An entire year of unlimited access costs $39 or less. I always maintain a link to their best sales on my website under the Books menu item. PIA allows you to connect to your choice of dozens of servers worldwide. I can choose California when I want to appear on the west coast or New York when I want to appear on the east. I can choose London in order to bypass restrictions while watching the BBC online or Toronto when I need to appear as a Canadian user. Your yearly access can be used on up to five devices simultaneously. My personal policy on VPNs is quite simple. I always use a

VPN on any device that I connect to the internet. This includes desktops, laptops, and cell phones.

To be transparent, I am a PIA affiliate. This basically means that they give me a few bucks each time I refer someone to their paid service from my website. However, I was a paid user of their product for two years prior to recommending their services. If you are interested in learning the reasons why I use PIA, please read my online PIA page at the following site.

https://privacy-training.com/pia.html

Summary

I believe that executing much of the content in this chapter is vital before proceeding through the rest of this book. The methods discussed here help protect you, your computers, and your internet connection. The browsers and extensions make online research easier and more efficient. The casual searcher may have no need for virtual machines or VPNs. However, the global security team that monitors threats from violent people cannot afford to go without these. Consider the following summary, in order of most applicable to specialized needs.

- Ensure that your computers are protected from known viruses and malware.
- Execute a system cleaner at least once weekly.
- Install a better browser such as Firefox or Chrome.
- Consider browser add-ons that will make your research more efficient.
- Connect via a VPN for a secure and masked internet connection.

CHAPTER TWO
BUSCADOR LINUX VIRTUAL MACHINE

Linux operating systems have been a part of my OSINT investigations and trainings for many years. They are lightweight, run on practically any hardware, cost nothing, and provide a level of security that cannot be obtained through traditional operating systems such as Microsoft Windows. During my training, I often demonstrate how I use Linux as a virtual machine or bootable USB device. In both scenarios, I can navigate to any malicious website, download every virus possible, and eliminate all traces of my activity by simply rebooting the system. Upon reboot, there are no viruses and everything works exactly as intended when the system was created.

This chapter presents ways that you can harden your security by using a Linux operating system during your investigations. Many years ago, this may have been intimidating to non-technical users. Today, implementing Linux into your investigations is extremely easy. This chapter is intentionally at the beginning of the book in order to better protect you and your investigations right away. Once we start exploring the world of online search techniques in Chapter Three, you will likely encounter malicious software or viruses at some point. If you investigate cyber criminals, this will be sooner rather than later. The malicious code will almost always target Windows machines. By choosing Linux as your host system, you greatly lessen the concern about infections. This chapter will vary from basic tutorials through advanced technologies. As stated in the introduction, some readers may want to skip portions of this chapter in favor of the traditional search techniques that begin in Chapter Three. I present material on Linux before online searching because I want you to have a safe and secure environment for your research, without the fear of harming your personal computer.

In 2015, I actively taught methods that would take a standard Linux system, such as Ubuntu or Mint, and install it to a USB device. This device could be inserted into a computer, booted, and a native Linux system would be present. When the machine was turned off, all history of that session was eliminated. This was a quick and easy way to conduct high-risk investigations while protecting the integrity of a personal computer. Unfortunately, this was slow, mostly due to the speed bottleneck of the USB device. It was a valid practice with good intentions, but not extremely applicable to most OSINT investigators. The previous edition of this book had an entire chapter devoted to creating these devices.

In late 2016, I was contacted by David Westcott. We knew each other through our OSINT work, and he asked if I was interested in creating a custom OSINT virtual machine. I had always considered this, but had concerns about my skills at hardening Linux systems and pushing out finished builds. David had worked on other public Linux releases, and was much more comfortable distributing custom systems. I began designing my dream OSINT build, sending him weekly requests, and he began taking the ideas and executing them within a test product. By

early 2017, the first version of our new operating system was released and titled Buscador (*Seeker* in Spanish).

This concept is not new. Many readers are likely familiar with Linux systems such as Kali. This is a custom Linux build that includes hundreds of security testing tools for the cyber security community. It is considered an all-in-one operating system that has everything pre-configured upon installation. We wanted that same experience for the OSINT community. Buscador was designed from the ground up with considerations for OSINT investigations. The web browsers are pre-configured with custom settings and extensions, and numerous OSINT software applications are already set-up to accept your search queries.

An important part of Buscador is the applications. On many Linux builds, launching software is not similar to traditional operating systems. While the software is installed, you must still launch a Terminal window and type the specific commands required. This can be very difficult and unforgiving. There are seldom point-and-click icons that launch similar to Windows. This has always created a barrier between the geeks and the norms. Either you know how to issue Linux commands or you do not. If you don't, then you never get to take advantage of the power of Linux and Python.

We wanted to eliminate that barrier. We wanted to make powerful Linux programs easily accessible to everyone. My initial thought was to create Bash scripts similar to batch files in Windows, but David came up with a much easier and more appropriate way. Every tool inside Buscador has its own icon in the dock, executes by clicking with a mouse, and walks the user through the menus. After collecting the required data, each program executes the proper commands behind the scenes and delivers the content directly to the user. We believe this to be unique in our community. Every person, at any skill level, can use Buscador as a virtual machine.

Virtual Machines

Virtual machines (VMs) conduct emulation of a particular computer system. They are computer operating systems on top of computer operating systems. Most commonly, a software program is executed within an operating system, and individual operating systems can launch within that program. Each virtual machine is independent from the other and the host operating system. The environment of one virtual machine has no impact on any others. Quite simply, it is a way to have numerous computers within your single computer.

Before installing Buscador, you must possess virtual machine software. There are several programs that allow you to create and execute virtual machines. Many of these are paid programs, such as VMWare. However, I will focus on VirtualBox in this section. VirtualBox is completely free and easy to operate. All methods presented here for VirtualBox can be replicated on VMWare, and I will identify some key differences when appropriate.

VirtualBox (virtualbox.org)

Volumes could be written about the features and abilities of VirtualBox. I will explain how to configure and maintain a virtual machine within the application. Installation instructions can be found at virtualbox.org and are very straightforward. You would install VirtualBox in the same manner that you would install any other application in Windows, Mac, or Linux. The default options during setup will be sufficient for our needs. The only requirement for VirtualBox to function is a computer that supports virtualization.

Any modern Apple product will work without any modification. Most mid-range and high-end Windows computers made within the past five years should have no problem, but may require you to enable virtualization support in the BIOS (Basic Input / Output System). Netbooks, older machines, and cheap low-end computers will likely give you problems. If you are in doubt about meeting this requirement, search for your model of computer followed by virtualization and you should find the answers. The rest of this section will assume that your computer meets this requirement. All instructions here apply to Windows, Mac, and Linux operating systems. After installing VirtualBox, you should install the VirtualBox Extension Pack located on the VirtualBox website.

Buscador Download

Buscador is completely free and can be found at **https://inteltechniques.com/buscador**. The user name for this software is **osint**, and the password is also **osint**. These can be changed later. This is a large file, usually over 3GB. I recommend storing the file that you download for future use in case you want to rebuild your system. On the download page, there are three options: VMWare, VirtualBox, and ISO. If you will be using VirtualBox (free), choose that file. If you have VMWare Workstation (Windows) or Fusion (Mac), then you will want the appropriate file. Overall, David and I use VMWare as it tends to offer a smoother experience. However, most users have VirtualBox with no complaints. I encourage you to test with VirtualBox, and upgrade to VMWare if you find yourself restricted by the limitations of VirtualBox.

Buscador VirtualBox Installation

- In the VirtualBox menu, click on File > Import Appliance
- Navigate to the OVA file that was downloaded (Buscador)
- Choose this file and select "Import"
- Before starting the new machine, highlight it and choose "Settings"
- Under General > Basic, rename this machine as desired (Buscador?)
- Under General > Advanced, change Shared Clipboard to Bi-Directional
- Under System > Motherboard, increase RAM to half of total system resources
- Under Display > Screen, increase the Video Memory to 128MB if available
- Under Storage, click the small "plus" in the lower left corner

- Click "Add Optical Drive" and "Leave Empty"
- Under Shared Folders, click the "plus" on the right
- Choose a local folder to store evidence and select "Auto-Mount"
- Under Audio, enable audio and select host machine's audio controller
- Click "OK" and launch the new machine
- Upon boot, log into the user "osint" with the password of osint
- In the VirtualBox Menu, select Devices > "Insert Guest Additions CD Image"
- Allow the image to be installed, and reboot upon completion
- Start the Terminal in the new VM and type sudo adduser osint vboxsf
- Provide the password as needed (osint)
- Reboot

You should now have access to the shared directory in order to save data to the host operating system (evidence). It can be found in the File Manager (Home), on the left column, titled "sf_" followed by the name of the folder to which it is connected. This shared folder will also be on your desktop for easy access. You can make the machine full-screen, copy and paste text to and from the image, and you are ready to begin using the applications.

Buscador VMWare Installation

- In the VMWare menu, select File > Import > Select OVA
- Select the location where the VM will be imported. Click "OK"
- Click "Retry" if the initial import fails
- Power on the VM and Login to the OS
- Install VMWare tools as appropriate for your version:
  VMWare Fusion: In the menu, select Virtual Machine > Install VMWare Tools
  VMWare Workstation: In the menu, select VM > Install VMWare Tools
  VMWare Player: In the menu, select Player > Manage > Install VMWare Tools
- Open (Double Click) the VMWare Tools CD mounted on the desktop
- Right-click the file that is similar to VMWare.xx.tar.gz and click "Extract to"
- Select the Desktop folder
- Open Terminal (Type 'No' to avoid update) and type cd Desktop/VMWare-tools-distrib
- Type sudo ./VMWare-install.pl and enter password (osint)
- Type Y when prompted about downloading from the Linux repository
- Accept all default values by striking the enter/return key at every prompt
- Reboot the VM
- Enable Shared Folders from the file menu: Settings > Options > Shared Folders
- Add a Shared Folder by selecting the desired folder on the host OS
- Create a shortcut to the folder on the desktop with the following Terminal command
  ln -s /mnt/hgfs/foldername/home/osint/Desktop/Shared_Folder

Updates

Upon creation of your new Buscador virtual machine, you should apply all system updates by clicking on the "Software Updater" in the lower left portion of the dock. This will update your core system and standard applications such as your browsers. Next, launch the Terminal (first black box icon in the dock). If prompted to update, type Y and striking the enter key. Finally, type "update_scripts" within this same Terminal window to apply any updates that we have made since the last stable release. You are now ready to preserve this machine's state.

Snapshots

A great feature of virtual machines is the use of Snapshots. These "frozen" moments in time allow you to revert to an original configuration or preserve an optimal setup. Most users install the virtual machine as detailed above, and then immediately create a snapshot of the unused environment. When your virtual machine eventually becomes contaminated with remnants of other investigations, or you accidentally remove or break a feature, you can simply revert to the previously created snapshot and eliminate the need to ever re-install. Consider how I use snapshots, as detailed below.

Upon creation of a new Buscador virtual machine, I apply all updates as mentioned previously. I then completely shut down the machine and open the Snapshots option with my virtual machine software. I create a new snapshot and title it "New Install". I then use this machine for a single investigation, and export all evidence to an external USB device, such as a flash drive. I then "restore" the New Install snapshot, and it overwrites any changes made during the previous investigation. Upon reboot, all history and evidence is eliminated. This ensures that I never contaminate one investigation with another. When there are substantial updates available for Buscador, I load the default configuration, and apply all updates. I then shut the machine down completely and delete the New Install snapshot, without saving it, and create a new snapshot titled New Install. This new snapshot possesses all of the updates and I repeat the investigation process. I usually delete and create a new snapshot weekly. The use of snapshots is very similar between VirtualBox and VMWare, but let's take a look at the minor differences.

VirtualBox use of Snapshots

- Completely shut down the Virtual Machine
- In the VirtualBox Menu, click on the Snapshots button in the upper right
- Click on the blue camera icon to "take a snapshot"
- Create a name and notes to remind you of the state of the machine, such as "New Install"
- Click OK

You can now use your virtual machine as normal. If you ever want to revert to the exact state of the machine that existed at the time of the snapshot, follow these instructions:

- Completely shut down the Virtual Machine
- In the VirtualBox Menu, click on the Snapshots button in the upper right
- Select the desired snapshot to apply
- Click on the blue camera icon with arrow to "restore snapshot"
- Click Restore

Optionally, if you ever want to remove a snapshot, simply use the icon with a red X. This will remove data files to eliminate wasted space, but you cannot restore to that image once removed. It will not impact the current machine state. Many users remove old, redundant snapshots after creating newer clean machines.

VMWare Use of Snapshots

- Completely shut down the Virtual Machine
- In the VMWare Menu, click on the Snapshots button in the upper right
- Click on the camera icon to "take" a snapshot
- Create a name and notes to remind you of the state of the machine, such as "New Install"
- Click Take

You can now use your virtual machine as normal. If you ever want to revert to the exact state of the machine that existed at the time of the snapshot, follow these instructions:

- Completely shut down the Virtual Machine
- In the VMWare Menu, click on the Snapshots button in the upper right
- Select the desired snapshot to apply
- Click on the camera icon with arrow to "restore" a snapshot
- Click Restore

Optionally, if you ever want to remove a snapshot, simply use the "delete" icon. David suggests to enable VMWare AutoProtect snapshots, set to daily, and limit the snapshot count to 3. After the maximum number of AutoProtect snapshots is reached, Workstation deletes the oldest AutoProtect snapshot each time a new AutoProtect snapshot is taken. This setting does not affect the number of manual snapshots that you can take and keep. AutoProtect snapshots are an easy way to always have a snapshot to revert to later. I do not use this technique as I want more manual control of my snapshots. The following steps will enable this feature.

- Select the virtual machine and select VM > Settings
- On the Options tab, select AutoProtect and select Enable AutoProtect
- Select the "Daily" interval between snapshots
- Select the maximum number of AutoProtect snapshots to retain (Recommended "3")
- Select OK to save your changes

If you ever want to preserve a specific state of Buscador, you can export an entire session. This may be important if preserving your work environment for court purposes. When I am conducting an investigation that may go to trial, or discovery of evidence will be required, I make an exact copy of the operating system used during the investigation. At the end of my work, I shut down the machine. I click on File and then Export within my virtual machine software and create a copy of the entire operating system exactly as it appeared at shutdown. This file can be imported later and examined. After successful export, I restore my clean New Install snapshot and I am ready for the next case. The exported file is added to my digital evidence on an external drive. I now know that I can defend any scrutiny by recreating the exact environment during the original examination.

Hopefully, you now have either VirtualBox or VMWare installed and Buscador imported as a virtual machine. Figure 2.01 displays the Buscador operating system with application dock to the left. Scrolling up and down while hovering over this dock presents all of the applications available. Clicking the nine small dots in the upper left opens a list of applications. This is redundant, but may be easier to access. Now it is time to play with the many applications inside Buscador. Note that the following instruction was created from Buscador version 1.2, which was released at the same time as this book. Each version may vary slightly, but the concepts should remain constant.

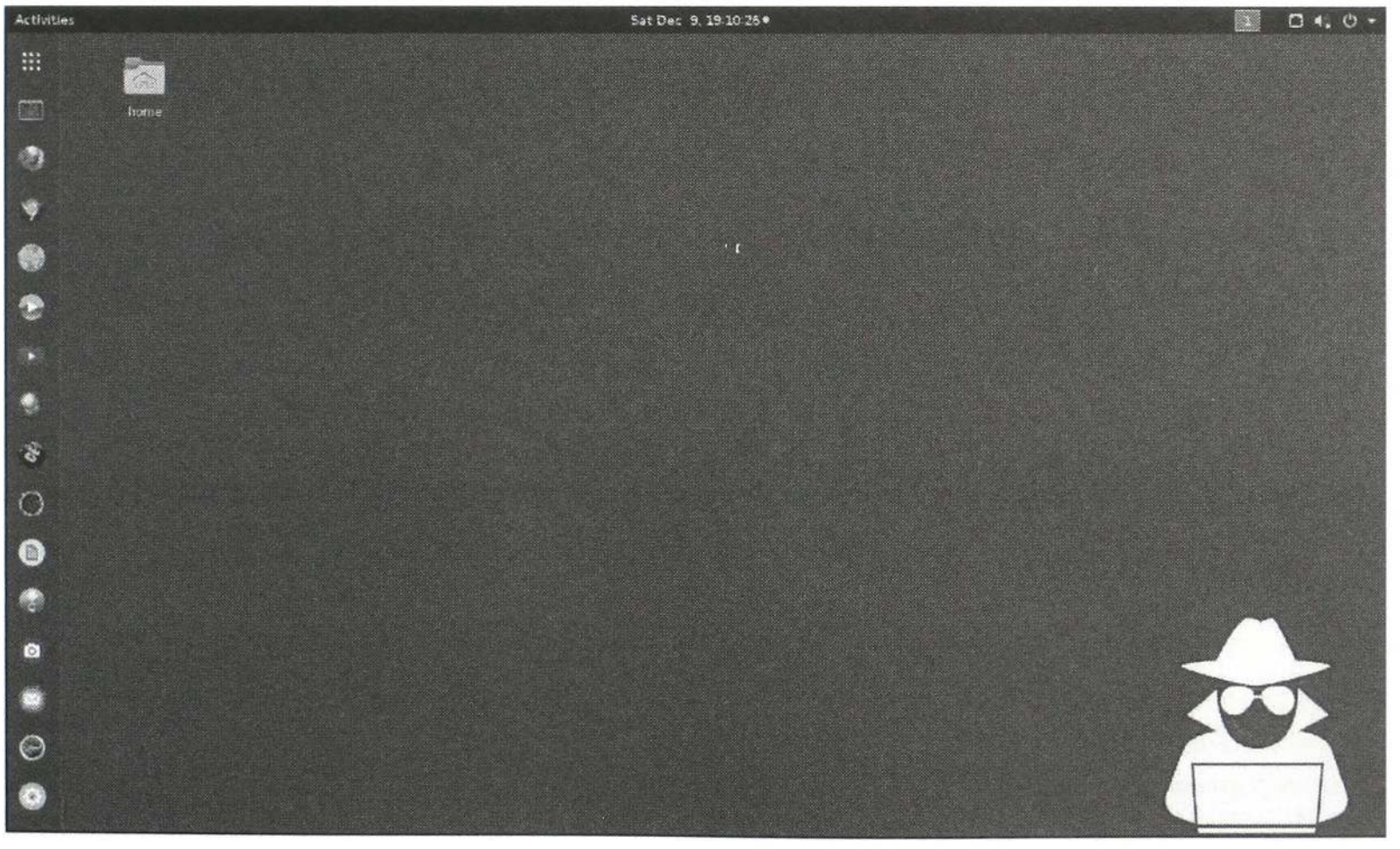


Figure 2.01: The Buscador virtual machine.

Web Browsers

In Chapter One, I explained how to harden the security of Firefox and introduced several add-ons that aid in OSINT investigations. Every piece of that instruction has been applied within Buscador. When you launch Firefox, you will see the configuration changes and extensions right away. You do not need to import any profiles or modify any settings. The Chrome option directly below Firefox is set to run in Incognito mode to prevent your internet history from being recorded. Again, several extensions have already been installed as mentioned previously. The bookmarklets have also been added. The third option is the Tor Browser. It is the exact version that was explained in Chapter One, and it is also ready to be used. Overall, the entire instruction of the previous chapter has been applied within Buscador. As we release updates, we will continue to apply changes to strengthen the virtual machine.

Video Utilities

This is the first application that should demonstrate the power of Buscador combined with the ease of a point-and-click system. This option executes an open-source video tool called FFmpeg. The icon launches a dialogue that firsts notifies you that "The next window will prompt for a target media file". After clicking OK, you are presented with a file manager window that is requesting that you choose a target video file. This could be a downloaded YouTube video, an unplayable surveillance video, or any other downloaded video content retrieved from any source. After you select the video, you are presented with a menu of the following options, with added explanations of the usage.

Play a video: This option will force FFmpeg to attempt to play any video file with multiple video codecs. This will often play videos that would not otherwise play using standard media players such as Windows Media Player and VLC. This will also play many surveillance videos without the need for a third-party program, and downloaded live streams that have no associated player.

Convert a video to MP4: This option simply converts any target video to a standard MP4 format. This is beneficial when the target video possesses a unique codec that prohibits playing universally. If the above option can play the video, this option can convert it so that any computer should be able to play it natively. The video that is created will be stored in the "Video Utilities" folder in the left menu of the "Home Folder" on the Desktop.

Extract video frames: This is likely the most used utility within this set of applications. After supplying a target video, this tool will extract the still images from the video. The result is a folder of uncompressed bitmap (bmp) image files depicting practically every frame of the video. This is beneficial for evidence when close examination of single frames is necessary. Figure 2.02 displays a result which created tens of thousands of still images. The created images will be within a new folder in the same "Video Utilities" folder mentioned previously, as seen in the example. This option can fill your virtual hard drive space quickly, so be sure to revert to a clean snapshot when finished.

Shorten a video (Low activity): I originally added this option within a Windows-based version of this utility for a specific purpose. I had a surveillance video over two hours in length. It did not use motion activation, and I needed to watch it to identify if anyone entered a room. Instead of sitting for two hours, I used FFmpeg to remove the frames without action. This version of the script removes single frames of a video that appear extremely similar to the frame preceding it. In other words, it takes out all of the frames which are the same (no action) and only leaves the frames where there is some type of activity. The result in my case was a two-minute video isolating only the portion with activity.

Shorten a video (High activity): This version of the script is identical to the previous with one exception. It is a bit more forgiving for videos with high activity. This might be outdoor surveillance of people walking in the distance or a video with a time counter printed within the feed. If the previous option does not remove enough content, this version will be a bit more aggressive.

Extract audio: This option extracts the raw audio file out of any video, converts it to a 320k MP3 file, and places it in the "Video Utilities" folder within the Home folder on the Desktop. I have used this to extract audio from video confessions and interviews, and it works well on any online videos downloaded from the internet.

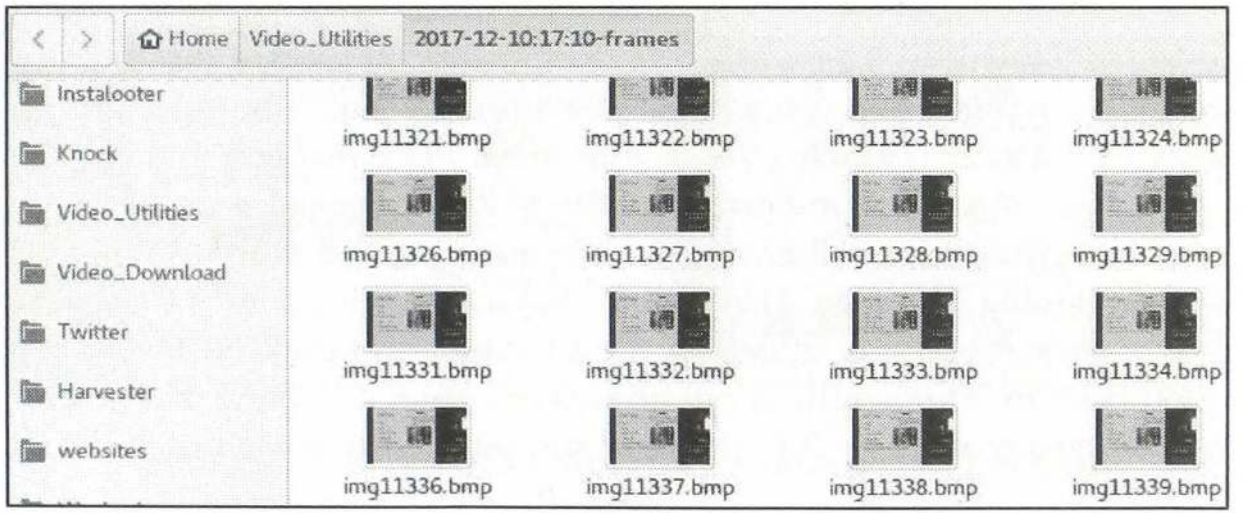


Figure 2.02: Still images created from a target video with the Video Utilities application.

Some advanced readers may wonder what is happening behind the scenes during these executions. When Buscador prompts you for the target video, it enters the supplied data path into a pre-configured terminal command. As of this writing, the basic commands are as follows. Note that (Video) refers to the actual file name of the video.

Play a video:
ffplay (Video)

Convert a video to MP4:
ffmpeg -i (Video) -vcodec mpeg4 -strict -2 $timestamp.mp4"

Extract video frames:
ffmpeg -y -i (Video) -an -r 10 **$timestamp-frames/img%03d.bmp**

Shorten a video (Low activity):
ffmpeg -i (Video) -strict -2 -vf "select=gt(scene\,0.003),setpts=N/(25*TB)" **$timestamp.mp4**

Shorten a video (High activity):
ffmpeg -i (Video) -strict -2 -vf "select=gt(scene\,0.005),setpts=N/(25*TB)" **$timestamp.mp4**

Extract audio:
ffmpeg -i (Video) -vn -ac 2 -ar 44100 -ab 320k -f mp3 $timestamp.mp3"

Video Download

This application is powered by a Python script called YouTube-DL. Python is a programming language that is very common in the OSINT and computer security communities. It is easier to learn than other programming options and globally universal across Linux operating systems. YouTube-DL is an extremely powerful utility that has been reserved for those that understand Python and command-line input. This will be the first application that we free from this restriction.

Click on the Video Download icon within the dock to launch a script that will prompt you to "Enter Video URL". Assume that you want to download the YouTube video located on the site at https://www.youtube.com/watch?v=6kBOCnOlwqI. Place that link in the prompt and click "OK". An animated screen will indicate that the video download is in progress. When the download completes, Buscador will launch the file manager, and display the default folder for video downloads (Home > Video_Download). If you ever need to navigate to this folder manually, double-click the Home folder on the Desktop, and click on the Video_Download folder in the left column. These options can be seen in Figure 2.03 (left). The downloaded video will be the highest quality available for download and the file name will match the video title.

Next, assume that we are looking for videos of Bob Ross teaching viewers how to paint with oils. After a search on YouTube of Bob Ross, I found the official Bob Ross channel located at https://www.youtube.com/user/BobRossInc. Clicking on the Videos option on that page navigated me to https://www.youtube.com/user/BobRossInc/videos. This page displays over 600 full episodes of his television program. I relaunched the Video Download option in the Buscador dock and supplied the URL. Buscador began to download all of the videos, one at a time, from that page. At completion, the File Manager opens and displays all of the files. Figure 2.03 (right) displays this process in action.

You may wonder what is happening behind the scenes during these downloads. When Buscador prompts you for the URL of the video or bulk video page, it enters the supplied data into a pre-configured terminal command. As of this writing, the basic command is as follows with an explanation after each.

youTube-dl (Video) -f 'best[ext!=webm]' -o /home/osint/Video_Download/"%(title)s.%(ext)s" –i

youtube-dl: The command to execute the script
(Video): The collected URL of the target page
-f 'best[ext!=webm]': Forces download of best available quality and ignores webm versions
-o /home/osint/Video_Download/: Specifies location of downloaded media
"%(title)s.%(ext)s": Titles the video file with the video name and extension
-i: Ignores any errors

Additional options to include in this command include the following.

--all-subs: Downloads any closed captioning subtitles associated with the video(s)
--all-formats: Downloads all versions of a video of any quality

While named YouTube-DL, this script works on most popular video websites. You should have no issues downloading individual or bulk videos from YouTube, Vimeo, LiveLeak, WSHH, and many others. The bulk download option has saved me numerous hours of manually downloading videos individually. I have yet to find any size or file number limitations. This utility is likely the most used program within Buscador, aside from web browsers.

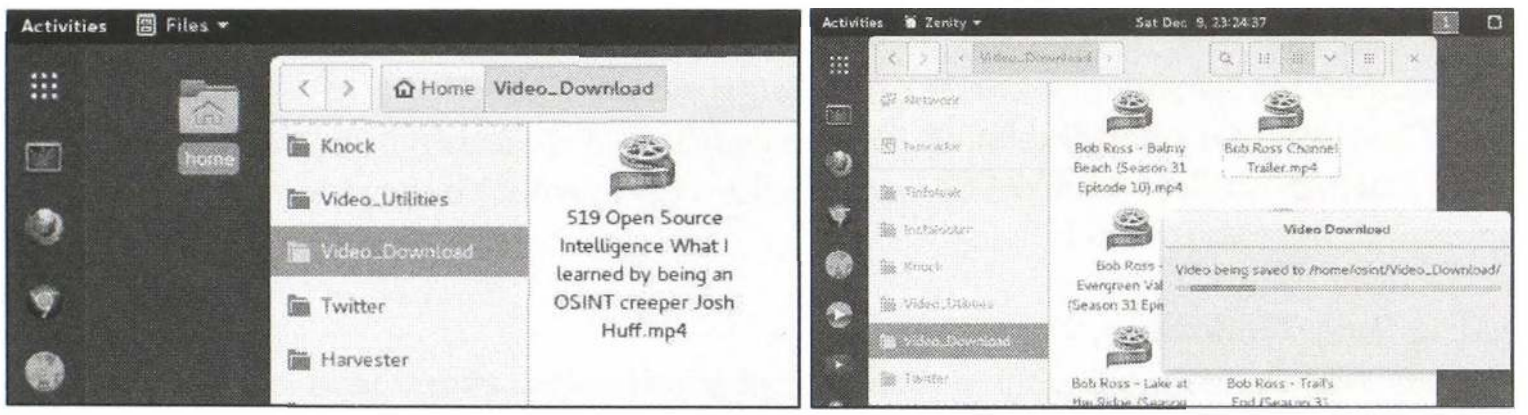


Figure 2.03: The Video Downloads location within Buscador.

Hopefully by now, you see the benefit of using a custom-built Linux operating system with pre-configured applications versus a command-line driven traditional system. I may have over-explained the previous options, but I wanted to give a sense of the goals of making this system user-friendly. The remaining options will be explained more briefly as not to overwhelm any readers with unnecessary details.

Metagoofil

In Chapter Nineteen you will read about a Windows program called FOCA that downloads documents and extracts the metadata from each. The importance of document metadata is also explained in that chapter. FOCA has been a staple of my investigations for many years, but Metagoofil was around before FOCA was created. Metagoofil is a command-line utility that

automates the collection of documents from a specific domain, and analyzes the metadata after retrieval. Clicking this option in the dock of Buscador will prompt you for a domain name such as IntelTechniques.com. You are then prompted for the maximum number of documents to retrieve. This will prevent hours of unintended data download from large websites. This entered data will generate a command to identify documents stored on that domain, and then download each into the Metagoofil directory in the Home folder.

As an example, I performed a search of cisco.com and limited my results to 100 documents. Metagoofil retrieved 100 various PDF, Word, and Excel files from the Cisco domain. It then created four text files that incorporate the metadata stored within these documents, with the following explanation of each.

Full.txt: This is the full metadata export created by Metagoofil. It includes all metadata available about all files downloaded. Below is an actual excerpt from the example conducted for this tutorial. I now know the real names and user names of two people associated with the usage of that document, dates and times of creation and modification, and the company that employs the individuals involved. These are valuable pieces that would otherwise go unnoticed.

File: SX_SY_EFSU_Compatibility_Matrix.xlsx
File Size: 43 kB
Creator: Chi Iong Ansjory (ansjory)
Last Modified By: Niranjini Gurusubramanian (ngurusub)
Create Date: 2011:06:16 20:08:56Z
Modify Date: 2017:11:07 16:37:24Z
Application: Microsoft Excel
Doc Security: None
Heading Pairs: Worksheets, 11
Company: Cisco
Shared Doc: No

Authors.csv: This file extracts only the author names of the people that created the downloaded documents. This gets you right to a set of names for immediate analysis.

Companies.csv: This file extracts only the company details stored within all of the documents. This will immediately identify associations with other organizations. The domains of these companies could then be used for an additional search through Metagoofil.

Modified.csv: This file extracts only the metadata that identifies individuals that modified the downloaded documents. Similar to the authors option, it can disclose people associated with the target organization.

At the conclusion of my demonstration, I possessed 100 documents from my target's website that would have taken hours to locate and extract. Furthermore, I have a list of dozens of

employee names and additional companies connected to my target. All data is saved in the Metagoofil directory in the Home folder of Buscador. While FOCA tends to have a more powerful interface, it does not compete with the simplicity of Metagoofil. For those that want the exact extraction command passed during execution, it is listed below. Note that (Domain) refers to the actual address while (#) refers to the number entered during execution.

python metagoofil.py -d (Domain) -t pdf,doc,xls,ppt,docx,xlsx,pptx -w -l ($) -n (#) -o file.txt

Harvester

This tool combines two powerful options, with occasional redundancy within them. The idea is to provide a domain name, such as IntelTechniques.com, and allow two scripts to collect any email addresses that are publicly associated with the domain. The program uses search engines to obtain the associated addresses. This is beneficial for locating individual email addresses when investigating a company. These addresses can be further researched using the methods discussed in Chapter Eight. Launching the dock icon will present an input box requesting a domain name. Supplying that data will then present two options of theHarvester and EmailHarvester. The first option is the original script created for this purpose, while the second is an "improved" script that attempts to dig deeper for results. In my experience, neither is always better than the other, so both should be checked. For the curious, the basic commands executed behind the scenes are listed below. Note that (Domain) refers to the actual target web address.

python theHarvester.py -d (domain) -b all -f theharvester_$harvest.html
python3 EmailHarvester.py -d (domain) -s emailharvester_$harvest.txt -e all –noprint

EyeWitness

While one of my favorite scripts within Buscador, I hear very little about its use from members of my training. This Python script automates the collection of screen captures from websites. Imagine the following scenario. You are investigating a long list of website addresses that were provided to you as a lead. Maybe they were websites visited from a suspect's computer; a list of social network profiles discovered during the execution of Recon-ng (Chapter Twenty-Two); or just a self-created list of URLs associated to your investigation. Manually visiting each site and using a screen capture tool can be overwhelming. Instead, let's automate the task.

Launching the EyeWitness icon from the Buscador dock will prompt you with a menu for selecting a single URL or multiple URLs. If choosing a single URL, you are allowed to type it directly into the menu and it fetches a screen capture. If you choose the multiple URL option, it displays "The next window will prompt for the text file that contains the target domains". Clicking OK will open a file manager window that will allow you to select a text file on your machine that contains the suspect domains. You may need to create this file if one was not automatically generated during a previous investigation technique. To do this, click on the nine small dots in the upper left portion of Buscador to launch the Applications screen, and click on

gedit to open a text editor. Provide a list of suspect websites in this file, one per line, and save it to your Desktop (or anywhere you prefer). In this example, my text file appeared as follows.

Inteltechniques.com
Computercrimeinfo.com
Privacy-training.com
Twitter.com/inteltechniques
Automatingosint.com
Instagram.com/tomhanks

Now that I have a text file with a list of domains, I can select this file within EyeWitness. EyeWitness visits each domain and collects a screen capture, saving each to the install directory. This folder will open at the completion of the capture. At the top, there will be a new folder titled similar to 01242018_190646. This title is a date and time stamp which indicates that it was created on 01/24/2018 at 19:06:46 local time. This is beneficial for later documentation. Inside this folder are two items of most interest. The first is the "screens" folder and the second is the "report.html" file. The screens folder contains png images that depict a full screen capture of each supplied website. Figure 2.04 displays partial results from this search. As you can see, some file sizes can get quite large due to the amount of content, especially with image-driven social networks. The report.html file is much more robust than the individual captured images. It combines all of the captures into a report, and includes metadata about each site. This can be saved or printed, and can always be found in the Buscador Home folder under EyeWitness. Next to the image of my website, the report revealed the following.

http://Inteltechniques.com Resolved to:107.180.46.189
Page Title:IntelTechniques.com | OSINT Training by Michael Bazzell
x-powered-by:ASP.NET server:Microsoft-IIS/8.5 x-powered-by-plesk:PleskWin
last-modified:Sun, 07 Jan 2018 01:36:14 GMT, date:Wed, 24 Jan 2018 19:06:49 GMT

This can be very beneficial when you have dozens, hundreds, or even thousands of domains of interest. I once converted a list of over 300 suspect Twitter accounts into a folder of screen captures of each. This allowed me to quickly identify which accounts were truly valuable to my investigation by simply viewing the evidence similar to photos in an album. The report was several hundred pages, but was generated in only a few minutes.



Figure 2.04: Results from an EyeWitness automated screen capture.

Subdomain Utilities

In Chapter Sixteen, I explain investigative techniques for domain names, and the importance of searching for subdomains. A subdomain is a dedicated address that is part of a larger domain, such as pics.inteltechniques.com. While inteltechniques.com might be a main landing page, pics.inteltechniques.com could possess hidden content not available within the main website. Programs within Buscador such as Knock, Sublist3r, and Aquatone request a domain name and then fetch subdomains associated with the target. This is a vital step for any investigation into a domain name.

As an example, I provided cnn.com to each of these options after clicking the appropriate icons in the dock. Knock located dozens of subdomains such as games.cnn.com, go.cnn.com, and community.cnn.com. Sublist3r found most of those and a few that Knock missed. Aquatone replicated most of our research, but missed a few known options and found some new subdomains, including IP addresses associated with each. All of these results are likely public pages that could have been found with a search engine, but I have found numerous hidden pages this way when researching criminal websites. As with all of these tools, any available options should be used.

Twitter Export

This tool was created as part of my training at Justin Seitz's python training course. It requests a Twitter user name after launching from the dock. It then downloads the photos posted to the target account and creates a spreadsheet of the account activity. A new folder titled the name of your Twitter target is placed in the Twitter folder inside the Home directory of Buscador. Inside this new folder is the following content.

Photos: This folder contains all images extracted from the account.

Twitter.csv: This file contains account post history in the following format, including date and time, message, and any links. These tweets were extracted from the account tomhanks.

Sun Dec 10 19:04:50 +0000 2017,What kitten lost this? Hanx. https://t.co/5WCoUiiEiv
Fri Dec 08 19:52:31 +0000 2017,Lost hat? Cold head! Hanx. https://t.co/OYpmJPU4v4
Sat Nov 25 18:40:10 +0000 2017,Lost in Disneyland but turned in to Lost & Found! Hanx.
Mon Nov 20 05:41:15 +0000 2017,A lost pair is hard to find. Hanx https://t.co/D7hqiPmNXj
Fri Oct 20 17:11:30 +0000 2017,Still best wait a bit. Hanx. https://t.co/DCR6elKZjS

Followers.txt: This file contains a list of all people that follow the target on Twitter. The format is the Twitter name of the user followed by the real name provided. From this example, a partial result appears as "jenIam8675309, Jennifer Farrell".

Friends.txt: Finally, this file displays the people that your target follows, in the same format as the previous option.

In Chapter Five, you will learn more advanced ways to extract additional information from your target Twitter account, but this automated process allows for quick download of data with very little input. I use this for every Twitter target I have because the text files are in standard CSV (comma separated value) format and can be easily imported into other software.

Tinfoleak

The previous script was made custom for this training. Tinfoleak is a publicly available Python script that reaches deeper into an individual's Twitter account for non-standard details such as posted content. The Twitter Options icon in the dock asks for a target Twitter user name. It then generates a custom report and opens the document within your web browser. Inside it, you will find details such account creation date, time zone, Twitter clients used, user mentions, and references to those accounts with which he communicates the most. With our demo of Tom Hanks, we now know that he uses the celebrity marketing company WhoSay for 93 % of his tweets, while the remainder are posted from a traditional web client on his computer. He does not post from a mobile device. This information was not available directly from his profile online. For those that want to replicate this option in the Terminal, below is the command.

python tinfoleak.py -i -s --conv --social --hashtags --mentions --meta --media d --geo -(user name)

Instalooter

This utility attempts to extract media from a target's Instagram profile. While screen captures of an account may suffice for initial investigations, you should consider downloading any images directly from the source when possible. Launching Instalooter from the dock prompts you to choose either a hashtag or a user handle. A hashtag would be a topic such as #osint, while a user handle would be the actual account name such as tomhanks. Supplying these details will begin the extraction of all images associated with the search. After providing tomhanks as my target, Instalooter quickly downloaded all 117 photos from his account and placed them in a folder titled tomhanks within the Instalooter folder in my Home directory. It also opens the target folder upon completion to immediately display the evidence.

This will be the final mention of Tom Hanks in this book. You may be wondering why I keep using him as an example. In the first edition of this book, I conducted demonstrations with Chris Hadnagy from Social-Engineer.org as a suspect. He seemed like a nice guy and I liked his work. We have since become friends and have collaborated many times. In later editions, I used Amber MacArthur as a demo. I followed her on TechTV and other channels at the time, and she was my first introduction to podcasts. Since then, we have co-authored a book and I have appeared on her radio show several times. I detect a pattern here, so I am shooting "Big". Tom, call me.

SpiderFoot (spiderfoot.net)

SpiderFoot is a reconnaissance tool that automatically queries over 100 public data sources to gather intelligence on IP addresses, domain names, email addresses, host names, and more. You simply specify the target you want to investigate, pick which modules to enable, and then SpiderFoot will collect data to build up an understanding of all the entities and how they relate to each other. Within Buscador, SpiderFoot appears as a simple button within the menu bar of Firefox and Chrome. However, it is actually a complex installation of Python scripts, databases, and a web server. Fortunately, Buscador users do not need to worry about installation and configuration. This build is ready to use the moment you click the SpiderFoot button.

After launching SpiderFoot from your browser of choice, you should be presented with a menu that displays an empty scan history. Clicking the New Scan button launches the selection menu. Here you can provide a "Scan Name" for your search and select the depth. Many users will accept the default "All" option, which can take a very long time to complete. Instead of jumping in and collecting every possible resource, let's take a look at more strategic options. In my first example, my Scan Name is "Suspect Website" and my seed target is "inteltechniques.com". The last section allows the following four levels of search, with my recommendations for each.

All: This runs everything and should be used when you must know all details possible. It may need to run for over an hour. Likely overkill for most investigations.

Footprint: This option identifies the target's network perimeter, associated identities, and other information that is obtained through a lot of web crawling and search engine use. This is a lengthy search, and the most useful for standard OSINT investigations.

Investigate: This conducts basic scanning and querying of sources that may have information about your target's maliciousness. This should be used on sites that deliver malware.

Passive: This option collects as much information as possible without touching the actual target site. This prevents your target from knowing you were looking.

In this example, I chose the Footprint option and executed a scan. It took 52 minutes to complete. While my website is fairly minimal, SpiderFoot obtained 925 pieces of information during its scan. Of that data, the following would have been of most interest to my investigation.

- It connected all of my websites including my personal sites and Twitter profile
- It extracted associates' full names from my online presence
- It displayed 393 unique URLs of which I have linked in the past on my sites
- It identified 87 internal web pages for further investigation
- It extracted metadata from images within my web pages
- It identified software used to create online files

- It accurately associated me with a third-party training company
- It identified all internal web pages that require a login
- It identified all internal web pages that possess fillable forms

The benefit of using this application versus manual identification through the methods that will be explained later is the automation. While the results were lengthy and impressive, this scan took very little effort from me. Additionally, the ability to export all of the findings to a simple CSV file makes this program twice as valuable. Once SpiderFoot locates new data, it attempts to use that content within additional scans. The more data it locates, the longer the process will need to run. I usually execute any lengthy scans before leaving the office, and let them run until completion.

As a general rule, never hit the "back" button or "refresh" option in your browser while within SpiderFoot. It may take you out further than you desire or kill a process. Instead, embrace the following built-in features while on various pages.

- Clicking the Scans option on any page displays all active and finished scans
- Clicking the name of a scan from this page presents the entire project
- Clicking Browse within a project displays hyperlinks to the data obtained
- Clicking Export Data from the Browse page generates a spreadsheet with all data
- Revisiting the Scans page refreshes all updated data
- Selecting and Deleting a scan removes it completely from your machine
- Scans can be stopped or rescanned from the Scans page

In my second example, I created a new scan titled "Suspect Email" and searched a target email address. Since SpiderFoot detects the type of input and only scans appropriate services, I selected the "All" option and executed the scan. It completed in two minutes. It identified several compromised database entries (Chapter Eight), and leaked details on Pastebin (Chapter Thirteen) as well as seventeen accounts at various social networks. That is impressive for two minutes of searching. I include SpiderFoot as part of every email address investigation I conduct.

I really can't say enough good things about this free resource. The attention to detail and constant expanding of options by the author deserves a lot of respect. If your investigation into a website or IP address allows the time required to complete a scan, it is well worth the wait.

Metadata Anonymisation Toolkit (MAT)

Similar to how we investigate images and documents posted to the internet by our target, others may investigate the files that we upload as part of our covert operations. Image metadata will be explained later in Chapter Fourteen, and can include the make, model, and serial number of a camera or location data of the capture. Document metadata can include your name, computer name, and network login details. If you ever plan to upload documents or images to the internet,

you may want to remove all of the metadata that could compromise your investigation. MAT can clean or "scrub" individual and bulk files. Launching this application from the dock presents an interface that allows selection of files (Add) and metadata removal from selected files (Clean). The current version supports most popular image, document, and audio formats.

LibreOffice

LibreOffice is a free and open source office suite, comprised of programs for word processing (Writer) and the creation and editing of spreadsheets (Math). While there are other programs within this free suite, we chose to only include the two that would be most vital to OSINT uses. The Writer program can be used within Buscador for creating reports while the Math program can be useful with opening files created by the previous utilities. While not direclty associated with LibreOffice, we also included a PDF viewer and VLC, which plays the majority of the online media content that you will likely encounter.

Standard Applications

Aside from the custom execution of command-line scripts, Buscador also possesses several standard applications available on many platforms. These will all be properly explained in Chapter Nineteen while discussing various Windows programs. The following are pre-configured and ready for use within Buscador.

Creepy: Social network geolocation to identify embedded location details
MediaInfo: Video metadata utility for displaying hidden video data
ExifTool: Photo metadata utility for displaying hidden photo data
HTTrack: Website cloning utility
Google Earth: Extended functionality over Google Maps
VeraCrypt: File and container encryption software
KeePassXC: Cross-platform password manager
BleachBit: Purges and deletes unnecessary files created after online usage

Bootable USB Devices

I believe that every online investigation computer should have a selection of safe operating systems aside from the traditional host Windows or Mac options. By now, you may have your Buscador virtual machine configured and you are ready to attack. There is one last option that should be discussed. As stated earlier, the previous edition of this book dedicated an entire chapter to bootable USB devices. Once Buscador was available, I transitioned all of my training toward that environment and abandoned the USB boot devices. We didn't even make a USB option available during the first release. Several people contacted me stating that they still wanted the option of booting an investigation computer into a safe Linux environment without launching their host operating system. Since options are always good to have, I have abbreviated and updated the previous USB instruction into the most applicable information for our needs.

The general premise of this method is to create a USB drive that can be used to boot an entire operating system from itself. It requires no access to the primary hard drive, and does not read or write data to it. It would work without a hard drive being present in the computer. It is completely isolated from your important data and it will not leak any usage details to your regular system. There are many scenarios that support the use of a bootable USB.

You may want an operating system that you know is always clean of any malicious software or viruses. Typical browsing behavior on a daily basis is likely to bring in some type of unwanted software onto your device, even when using a virtual machine with Buscador. A USB device can contain a complete operating system that has never been used for any private activity. There is no contamination from previous use. This section will only focus on creating bootable USB devices that do not store any details, also referred to as stateless operating systems. We will walk through two options.

Buscador Linux: This option allows you to create a bootable USB device that will work on most Windows and Mac computers. It boots into the Buscador operating system from the USB drive. This does not have any impact on the operating system on the internal hard drive of the computer. Any changes that you make while in the USB operating system will NOT be preserved. This is referred to as stateless, since the state of the system is not maintained between uses. This provides a clean operating system on each use with no contamination from previous sessions. This is ideal in situations where you do not want to save any data or history. It is the simplest option for users with extremely sensitive investigations. It contains all of the tools otherwise found in the Buscador virtual machine.

Ubuntu Linux: This is designed for users that either have compatibility issues with the Buscador build or simply want a simpler option. This system will not possess any of the custom tools and will be very basic. However, it also will not have the unique footprint of a custom operating system that could jeopardize extremely sensitive investigations. Users can blend-in better with this system and not have an immediate appearance as an online investigator.

Before proceeding, it is important to note that this is the most technical portion of the chapter. You do not need to replicate any of these options to use Buscador as a virtual machine (which I recommend). Before using bootable USB devices, I encourage you to consider applying full disk encryption on the primary hard drive of the computer that you will be using. If using Windows, it is called BitLocker. If using Mac, it is called FileVault. The overall topic of encryption exceeds the scope of this book; however, I will sum up the main benefits for our needs. Encrypting the data on your primary hard drive prevents the new USB operating system from having access to that data. When booting to the new USB device, it will not have the ability to decrypt the primary device. This keeps your primary data secure. It also prevents malicious software such as viruses from spreading to the old drive.

Next, you need to choose the proper USB flash storage device. I highly recommend only using USB 3.0 devices and ports. If your computer does not have a USB 3.0 port, these methods might

be slow and unusable. Most laptops created in the past five years have this feature. The benefit of USB 3.0 versus 2.0 is speed. The 3.0 drives can be read at over ten times the rate of the 2.0 devices. Many Windows laptops designate USB 3.0 ports with a blue interior. Mac computers do not do this. The correct USB flash storage drive is as important as the appropriate port. While USB 3.0 is a standard, all drives do not function at the same speeds. You will find very cheap devices that are technically 3.0 drives, but barely function above 2.0 speeds. You will also find extremely expensive drives that operate at nearly the same speeds as internal hard drives. Your situation and budget will determine the most appropriate drive. All of my testing for this section was completed using Sandisk Ultra Fit USB 3.0 drives.

Creating the Drives

Now that you have your USB 3.0 drive and port ready to go, you need to create a bootable device. During these instructions, you will be asked to download files that contain the entire operating system within them. You will usually have an option of obtaining either a 32-bit or 64-bit operating system. Ultimately, your needs will vary based on your hardware. If you are using Apple products, such as a MacBook Pro or MacBook Air, you must always download the 64-bit option. If you are on a personal computer running Windows 8 or 10, you are also likely using a 64-bit processor. If you are using an older machine, or low-cost netbook, you might have a 32-bit processor.

Overall, I always recommend trying the 64-bit option first. If it fails, you may need to repeat with the 32-bit version. Researching your model of computer on Google may save you a future headache. Note that Buscador is only available as a 64-bit option. If you have a 32-bit machine, you can only use Ubuntu or another Linux option.

Windows Computers

If you are strictly a Windows user, you will find creation of bootable USB devices extremely easy. The following instructions will download an "ISO" file, install it to a USB drive, and configure the system to allow you to boot from the device. The steps will install either the Buscador or standard Ubuntu operating systems, but any other Linux install file could be used. The created bootable operating system stores absolutely no data about your activity. When you shut the system down, it forgets everything. This is a way to be sure that the operating systems you are using are free from malicious files. You will need a Windows computer to create this drive. From this Windows system, conduct the following steps in order.

- If installing Buscador, navigate to inteltechniques.com/buscador and download the ISO
- If installing Ubuntu, navigate to Ubuntu.com/download and obtain the appropriate ISO
- Insert your desired USB drive that will be overwritten
- Download the Rufus application at https://rufus.akeo.ie/
- Execute the program and choose your USB drive (choose carefully!)

- Choose a partition scheme of "MBR … for BIOS or UEFI Computers"
- Click the button similar to a CD and choose the appropriate ISO file
- Click "Start" and allow the process to complete

The result is a drive that will boot to the selected operating system. Your USB drive is now ready to be used as a boot drive. You will need to ensure that the BIOS of your Windows computer is set up to boot from USB before the internal drive. On most computers, hitting either the F2, F10, F12, DEL, or ESC buttons repeatedly when you first turn on your computer will present this menu. While each computer is unique, a search on Google for "Enter BIOS (your computer model)" will present the answers you need. While this boot drive should work on an Apple computer, we will now create an option more appropriate for them. I have found this method to be the most reliable way to make a bootable USB device, and it should work with practically any Linux Live ISO image such as TAILS (as discussed in *The Complete Privacy & Security Desk Reference, Volume I: Digital*).

Apple Computers

In the steps of the previous instruction, the option was chosen to create a USB drive that would boot to either a BIOS or UEFI computer. These terms refer to the way that the computer starts. BIOS is traditionally associated with Windows while UEFI is associated with Apple. That option allows the created drive to work on either a Windows or Apple computer, but it requires a Windows operating system to make the drive. The following steps will create a Linux USB boot drive from an Apple computer without the need for Windows.

- Navigate to https://etcher.io and download and install the Etcher software
- Launch and click the Select Image option. Select the ISO file that you downloaded
- Choose the destination USB drive that you would like to overwrite (choose carefully!)
- Click the Flash option

The result will be a USB boot device that will load either Buscador or Ubuntu natively on any modern Apple computer. The operating system should appear exactly as the version created previously with Rufus on Windows. The biggest difference is that this version will only work on Apple devices and not Windows. Insert this drive into a powered-down Mac and turn it on. Hold down the Option key while booting, and you should see a new option. It will appear as an orange icon at the far right of any other options. Choosing this loads Linux outside of your native Apple operating system.

Overall, I have found booting USB drives to be much more reliable on Apple computers versus Windows machines. I have also found the hardware of all modern Apple devices to be supported with most Linux builds. I personally use either an Apple MacBook Pro or dedicated Linux laptop for all of my investigations, testing, and training. While more expensive than most Windows systems, I believe the premium will pay off in the end. Users tend to get longer life out of the

hardware and less "software creep" that tends to make older Windows machines feel aged. If you plan on conducting numerous OSINT investigations or plan to use bootable USB drives consistently, I highly recommend Apple laptops or Linux builds from System76.

Whether you chose the Windows or Apple route, you should now possess a USB drive that contains an entire operating system within it. The benefits of these drives are not limited to OSINT related needs. The security of these systems is likely much stronger than any other option. Not only are you operating within a Linux environment which has very few viruses or vulnerabilities, but you are always using a new install. Since these devices do not remember any of your activity, rebooting them is equivalent to using a brand-new operating system. If you purposely infected the USB drive with any virus, it would be eliminated with a simple reboot.

Investigators can use virtual machines and bootable USB devices to establish credibility. I have testified on numerous occasions about my practices to secure evidence during computer forensic cases. With these methods, you can always state that your environment was free of any malicious software. The session did not contain any evidence or history from any other cases. Furthermore, you can honestly state that the entire session did not rely on any licensing or software requirements. Any person questioning the validity of your setup could easily replicate the operation for independent testing. By using open source solutions as this, you are presenting transparency into your investigation. By not using proprietary software, you are eliminating a debate that many attorneys attempt to pursue. An opposing party cannot inquire as to possible software licensing violations as an excuse to suppress evidence. This entire setup is 100% free.

CHAPTER THREE
SEARCH ENGINES

The first stop for many researchers will be a popular search engine. The two big players in the United States are Google and Bing. This chapter will go into great detail about the advanced ways to use both and others. Most of these techniques can apply to any search engine, but many examples will be specific for these two.

Google (google.com)

There are entire books dedicated to Google searching and Google hacking. Most of these focus on penetration testing and securing computer networks. These are full of great information, but are often overkill for the investigator looking for quick personal information. A few simple rules can help locate more accurate data. No book in existence will replace practicing these techniques in a live web browser. When searching, you cannot break anything. Play around and get familiar with the advanced options.

Quotation Marks

Placing a target name inside of quotation marks will make a huge difference in a quick first look for information. If I conducted a search for my name without quotes, the result is 47,300 pages that include the words "Michael" and "Bazzell". These pages do not necessarily have these words right next to each other. The word "Michael" could be next to another person's name, while "Bazzell" could be next to yet another person's name. These results can provide inaccurate information. They may include a reference to "Michael Santo" and "Barry Bazzell", but not my name. Since technically the words "Michael" and "Bazzell" appear on the page, you are stuck with the result in your list. In order to prevent this, you should always use quotes around the name of your target. Searching for the term "Michael Bazzell", including the quotes, reduces the search results to 8,770.

Each of these pages will contain the words "Michael" and "Bazzell" right next to each other. While Google and other search engines have technology in place to search related names, this is not always perfect, and does not apply to searches with quotes. For example, the search for "Michael Bazzell", without quotes, located pages that reference Mike Bazzell (instead of Michael). This same search with quotes did not locate these results. Placing quotes around any search terms tells Google to search exactly what you tell it to search. If your target's name is "Michael", you may want to consider an additional search for "Mike". If a quoted search returns nothing, or few results, you should remove the quotes and search again.

This search technique can be vital when searching email addresses or user names. When searching the email address of "Michael@inteltechniques.com", without quotes, I receive 8,070 results.

When I search "Michael@inteltechniques.com" with quotes, I receive only four results that actually contain that email address. When your quoted search, such as "Michael Bazzell", returns too many results, you should add to your search. When I add the term "police" after my name, the results reduce from 8,770 to 1,870. These results all contain pages that have the words "Michael" and "Bazzell" next to each other, and include the word "police" somewhere on the page. While all of these results may not be about me, the majority will be and can easily be digested. Adding the occupation, residence city, general interest, or college of the target may help eliminate unrelated results.

Search Operators

Most search engines allow the use of commands within the search field. These commands are not actually part of the search terms and are referred to as operators. There are two parts to most operator searches, and each are separated by a colon (:). To the left of the colon is the type of operator, such as site (website) or ext (file extension). To the right is the rule for the operator, such as the target domain or file type. The following will explain each operator and the most appropriate uses.

Site Operator

Google, and other search engines, allow the use of operators within the search string. An operator is text that is added to the search, which performs a function. My favorite operator is the "site:" function. This operator provides two benefits to the search results. First, it will only provide results of pages located on a specific domain. Second, it will provide all of the results containing the search terms on that domain. I will use my name again for a demonstration. I conducted a search of "Michael Bazzell" on Google. One of the results is a link to the website forbes.com. This search result is one of multiple pages on that domain that includes a reference to me. However, this search only displayed one of the many pages on that domain that possessed my name within them. If you want to view every page on a specific domain that includes your target of interest, the site operator is required. Next, I conducted the following exact search.

Site:forbes.com "Michael Bazzell"

The result was all seven pages on forbes.com that include my name within the content. This technique can be applied to any domain. This includes social networks, blogs, and any other website that is indexed by search engines.

Another simple way to use this technique is to locate every page that is part of a specific domain. A search query of site:inteltechniques.com displays all 560 pages that are publicly available on my personal website. This can be a great way to review all of the content of a target's personal website without attempting to navigate the actual site. It is very easy to miss content by clicking around within a website. With this technique, you should see all of the pages in a format that is easy to digest. Also, some of the pages on a website that the author may consider "private" may actually

be public if he or she ever linked to them from a public page. Once Google has indexed the page, we can view the content using the "site" operator.

Real World Application: While conducting private background checks, I consistently use the site operator. I was once supplied a cellular telephone number of an applicant for a very public position. A search on craigslist.org of this number revealed no results. A Google search of "site:craigslist.org" and the phone number revealed archived expired posts by the applicant promoting himself as a male prostitute. More methods of Craigslist searching are detailed later in this book. Additionally, using a search such as "site:amazon.com" and the target name can reveal interesting information. A recent background check of an applicant that signed an affidavit declaring no previous drug or alcohol dependencies produced some damaging results. The search provided user submitted reviews that he had left on Amazon in reference to books that he had purchased that assisted him with his continual addiction to controlled substances. Again, this result may have appeared somewhere in the numerous general search results of the target; however, the site operator directed me exactly where I needed to look.

File Type Operator

Another operator that works with both Google and Bing is the file type filter. It allows you to filter any search results by a single file type extension. While Google allows this operator to be shortened to "ext", Bing does not. Therefore, I will use the original "filetype" operator in my search examples. Consider the following search attempting to locate PowerPoint presentation files associated with the company Cisco.

"Cisco" "PowerPoint"

The result is 909,000 websites that include the words Cisco and PowerPoint in the content. However, these are not all actual PowerPoint documents. The following search refines our example for accuracy.

"Cisco" filetype:ppt

The result is 28,300 Microsoft PowerPoint presentations that contain Cisco within the content. This search only located the older PowerPoint format of PPT, but not newer files that may have the PPTX extension. Therefore, the following two searches would be more thorough.

"Cisco" filetype:ppt
"Cisco" filetype:pptx

The second search provided an additional 12,000 files. This brings our total to over 40,000 PowerPoint files which is overwhelming. I will begin to further filter my results in order to focus on the most relevant content for my research. The following search will display only newer PowerPoint files that contain the exact phrase Cisco Confidential within the content of the slides.

"Cisco Confidential" filetype:pptx

The result is exactly 976 PowerPoint files of interest. There are many uses for this technique. A search of filetype:doc "resume" "target name" often provides resumes created by the target which can include cellular telephone numbers, personal addresses, work history, education information, references, and other personal information that would never be intentionally posted to the internet. The "filetype" operator can identify any file by the file type within any website. This can be combined with the "site" operator to find all files of any type on a single domain. By conducting the following searches, I was able to find several documents stored on the website irongeek.com.

site:irongeek.com filetype:pdf
site:irongeek.com filetype:ppt
site:irongeek.com filetype:pptx

Previously, Google and Bing indexed media files by type, such as MP3, MP4, AVI, and others. Due to abuse of pirated content, this no longer works. I have found the following extensions to be indexed and provide valuable results.

7Z: Compressed File
BMP: Bitmap Image
DOC: Microsoft Word
DOCX: Microsoft Word
DWF: Autodesk
GIF: Animated Image
HTM: Web Page
HTML: Web Page
JPG: Image
JPEG: Image
KML: Google Earth
KMZ: Google Earth
ODP: OpenOffice Presentation
ODS: OpenOffice Spreadsheet
ODT: OpenOffice Text
PDF: Adobe Acrobat
PNG: Image
PPT: Microsoft PowerPoint
PPTX: Microsoft PowerPoint
RAR: Compressed File
RTF: Rich Text Format
TXT: Text File
XLS: Microsoft Excel
XLSX: Microsoft Excel
ZIP: Compressed File

Hyphen (-)

The search operators mentioned previously are filters to include specific data. Instead, you may want to exclude some content from appearing within results. The hyphen (-) tells most search engines and social networks to exclude the text immediately following from any results. It is important to never include a space between the hyphen and filtered text. The following searches were conducted on my own name with the inclusion of excluded text. Following each search is the number of results returned by Google.

"Michael Bazzell" 8,770
"Michael Bazzell" -police 7,670
"Michael Bazzell" -police -FBI 7,000
"Michael Bazzell" -police -FBI -osint 6,010
"Michael Bazzell" -police -FBI -osint -books 4,320
"Michael Bazzell" -police -FBI -osint -books -open -source 404
"Michael Bazzell" -police -FBI -osint -books -open -source -"mr. robot" 9

The final search eliminated any results that included any of the restricted words. The nine pages that were remaining referenced other people with my name. My goal in search filters is to dwindle the total results to a manageable amount. When you are overwhelmed with search results, slowly add exclusions to make an impact on the amount of data to analyze.

InURL Operator

We can also specify operators that will focus only on the data within the URL or address of the website. Previously, the operators discussed applied to the content within the web page. My favorite search using this technique is to find File Transfer Protocol (FTP) servers that allow anonymous connections. The following search would identify any FTP servers that possess PDF files that contain the term OSINT within the file.

inurl:ftp -inurl:(http|https) filetype:pdf "osint"

The following will dissect how and why this search worked.

inurl:ftp – Instructs Google to only display addresses that contain "ftp" in the URL.

-inurl:(http|https) – Instructs Google to ignore any addresses that contain either http or https in the URL. The separator is the pipe symbol (|) located above the backslash key. It tells Google "OR". This would make sure that we excluded any standard web pages.

filetype:pdf – Instructs Google to only display PDF documents.

"osint" – Instructs Google to mandate that the exact term osint is within the content of the results.

Obviously, this operator could also be used to locate standard web pages, documents, and files. The following search displays only blog posts from computercrimeinfo.com that exist within a folder titled "wp" (WordPress).

inurl:/wp/ site:computercrimeinfo.com

InTitle Operator

Similar to InURL, the "InTitle" operator will filter web pages by details other than the actual content of the page. This filter will only present web pages that have specific content within the title of the page. Practically every web page on the internet has an official title for the page. This is often included within the source code of the page and may not appear anywhere within the content. Most webmasters carefully create a title that will be best indexed by search engines. If you conduct a search for "osint video training" on Google, you will receive 115,000 results. However, the following search will filter those to 863. These only include web pages that had the search terms within the limited space of a page title.

intitle:"osint video training"

Note that the use of quotation marks prevents the query from searching "video training" within websites titled "osint". The quotes force the search of pages specifically titled "osint video training". You can add "all" to this search to force all listed words to appear in any order. The following would find any sites that have the words osint, video, and training within the title, regardless of the order.

allintitle:training osint video

An interesting way to use this search technique is while searching for online folders. We often focus on finding websites or files or interest, but we tend to ignore the presence of online folders full of content related to our search. As an example, I conducted the following search on Google.

intitle:index.of OSINT

The results contain online folders that usually do not have typical website files within the folders. The first three results of this search identified the following publicly available online data folders. Each possess dozens of documents and other files related to our search term of OSINT. One provides a folder structure that allows access to an entire web server of content. Notice that none of these results points to a specific page, but all open a folder view of the data present.

http://cyberwar.nl/d/
http://bitsavers.trailing-edge.com/pdf/
http://conference.hitb.org/hitbsecconf2013kul/materials/

OR Operator

You may have search terms that are not definitive. You may have a target that has a unique last name that is often misspelled. The "OR" (uppercase) operator returns pages that have just A, just B, or both A and B. Consider the following examples which include the number of results each.

"Michael Bazzell" OSINT 1,390
"Mike Bazzell" OSINT 794
"Michael Bazzell" OR "Mike Bazzell" OSINT 1,800
"Michael Bazell" OR "Mike Bazell" OSINT 74
"Michael Bazzel" OR "Mike Bazzel" OSINT 75

Asterisk Operator (*)

The asterisk (*) represents one or more words to Google and is considered a wild card. Google treats the * as a placeholder for a word or words within a search string. For example, "osint * training" tells Google to find pages containing a phrase that starts with "osint" followed by one or more words, followed by "training". Phrases that fit this search include: "osint video training" and "osint live classroom training".

Range Operator (..)

The "Range Operator" tells Google to search between two identifiers. These could be sequential numbers or years. As an example, OSINT Training 2015..2018 would result in pages that include the terms OSINT and training, and also include any number between 2015 and 2018. I have used this to filter results for online news articles that include a commenting system where readers can express their views. The following search identifies websites that contain information about Bonnie Woodward, a missing person, and between 1 and 999 comments within the page.

"bonnie woodward" "1..999 comments"

Related Operator

This option has been proven very useful over the past year. It collects a domain, and attempts to provide online content related to that address. As an example, I conducted a search on Google with the following syntax.

related:inteltechniques.com.

The results included no references to that domain, but did associate it with my other websites, my Twitter page, my Blackhat courses, and my book on Amazon. In my investigations, this has translated a person's personal website into several social networks and friends' websites.

Google Search Tools

There is a text bar at the top of every Google search result page. This allows for searching the current search terms within other Google services such as Images, Maps, Shopping, Videos, and others. The last option on this bar is the "Search Tools" link. Clicking this link will present a new row of options directly below. This will give you new filters to help you focus only on the desired

results. The filters will vary for each type of Google search. Figure 3.01 displays the standard search tools with the time menu expanded.

The "Any time" drop-down menu will allow you to choose the time range of visible search results. The default is set to "Any time" which will not filter any results. Selecting "Past hour" will only display results that have been indexed within the hour. The other options for day, week, month, and year work the same way. The last option is "Custom range". This will present a pop-up window that will allow you to specify the exact range of dates that you want searched. This can be helpful when you want to analyze online content posted within a known time.

Real World Application: Whenever I was assigned a missing person case, I immediately searched the internet. By the time that the case is assigned, many media websites had reported on the incident and social networks were full of sympathetic comments toward the family. In order to avoid this traffic, I set the search tools to only show results up to the date of disappearance. I could then focus on the online content posted about the victim before the disappearance was public. This often led to more relevant suspect leads.

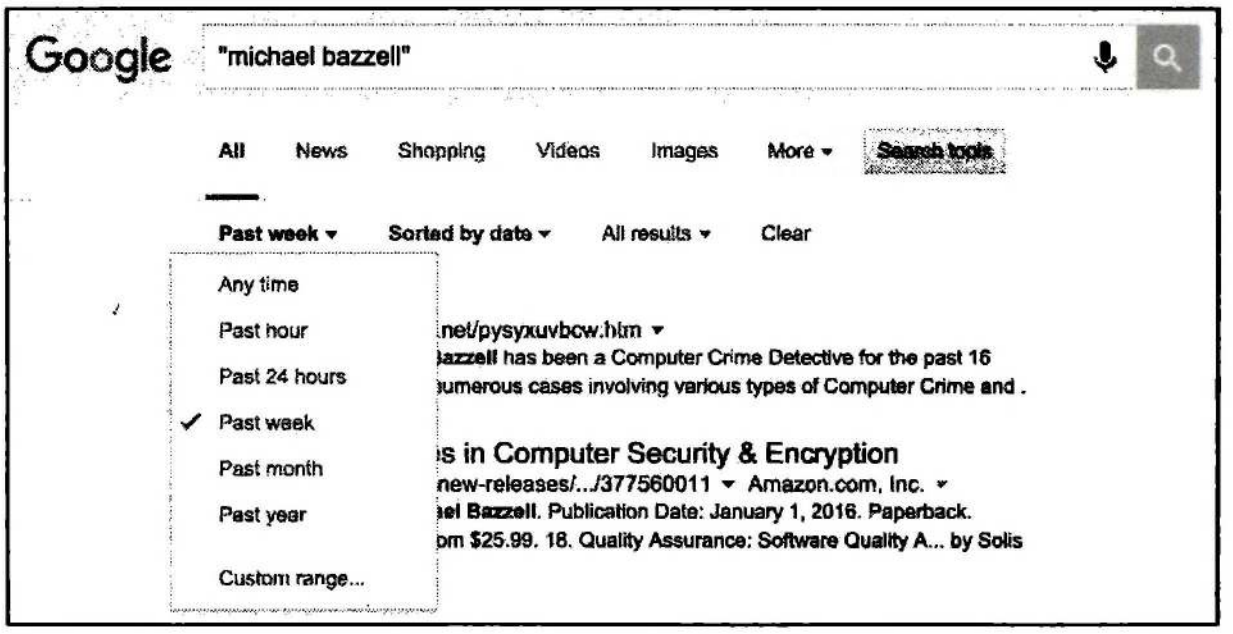


Figure 3.01: A Google Search Tools menu.

Dated Results

Google can be very sporadic when it comes to supplying date information within search results. Sometimes you will see the date that a search result was added to the Google index and sometimes you will not. This can be frustrating when you desire this information in order to identify relevant results. There is a fairly unknown technique that will force Google to always show you the date of each search result.

When you modify the "Any Time" option under the Search Tools menu, you will always see a date next to each result. If you are only searching for recent information, this solves the issue.

However, if you are conducting a standard search without a specific date reference, the dates next to each result are missing. To remedy this, you can conduct a specific search that includes any results indexed between January 1, 1 BC and "today". The appropriate way to do this is to add "&tbs=cdr:1,cd_min:1/1/0" at the end of any standard Google search. Figure 3.02 (top) displays the results of a standard search for the terms OSINT Tools. The exact URL of the search was "google.com/?#q=osint+tools". Notice that the result does not include a date next to the item. Figure 3.02 (bottom) displays the results of this same search with the specific data added at the end. The exact URL of this search was the following address.

google.com/?#q=osint+tools&tbs=cdr:1,cd_min:1/1/0

Notice that the result now has the date when the content was indexed by Google. You can also now sort these results by date in order to locate the most recent information. The search tools menu also offers an "All results" menu that will allow you to choose to see "all results" or "Verbatim". The All Results will conduct a standard Google search. The Verbatim option searches exactly what you typed. One benefit of the Verbatim option is that Google will often present more results than the standard search. It digs a little deeper and gives additional results based on the exact terms you provided.

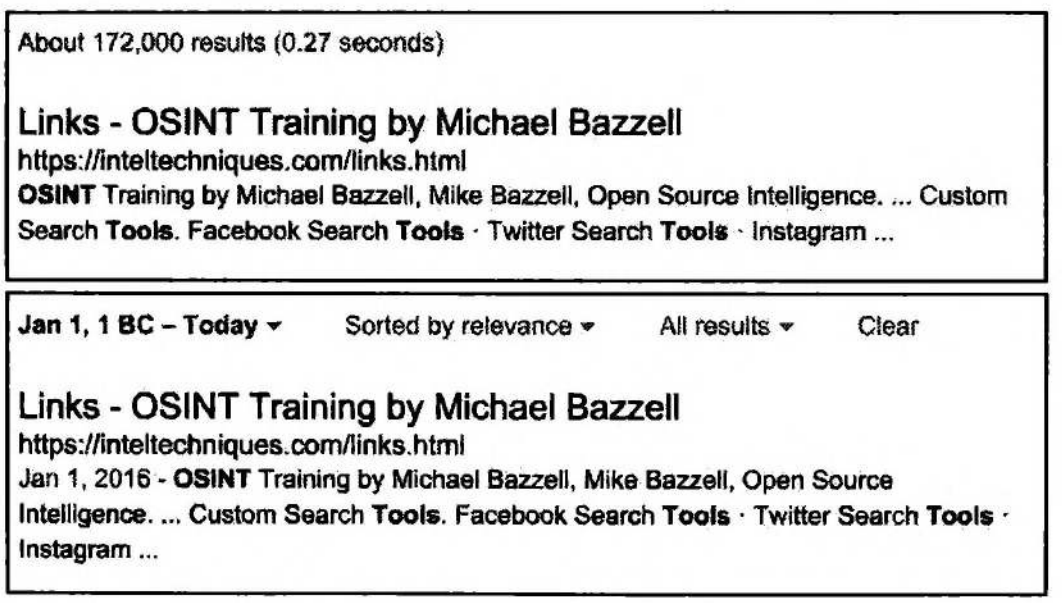


Figure 3.02: Results without (top) and with (bottom) date injection.

Google Custom Search Engines (google.com/cse)

Now that you are ready to unleash the power of Google, you may want to consider creating your own custom search engines. Google allows you to specify the exact type of searches that you want to conduct, and then create an individual search engine just for your needs. Many specialty websites that claim to search only social network content are simply using a custom engine from Google. For our first example, we will create a basic custom search engine that only searches two specific websites.

After you log into a Google account, navigate to the website listed above. If you have never created an engine, you will be prompted to create your first. Enter the first website that you want to search. In my example, I will search inteltechniques.com. As you enter any website to search, Google will automatically create another field to enter an additional website. The second website that I will search is computercrimeinfo.com. Provide a name for your custom engine and select "Create". You now have a custom search engine. You can either embed this search engine into a website or view the public URL to access it from within any web browser.

This basic functionality can be quite powerful. It is the method behind my custom Pastebin search engine discussed in Chapter Thirteen. In that example, I created a custom search engine that scoured 101 specific websites in order to retrieve complete information about specific topics. This is only the first layer of a Google custom search engine. Google offers an additional element to its custom engines. This new layer, labeled Refinements, allows you to specify multiple actions within one custom search engine. The best way to explain this is to offer two unique examples.

For the first example, I want to create a custom search engine that will allow us to search several social networks. Additionally, we will isolate the results from each network across several tabs at the top of our search results. The first step will be to create a new custom search engine in the same way that we did previously. Instead of specifying the two websites mentioned earlier, we will identify the websites to be searched as the following.

Facebook.com
Twitter.com
Instagram.com
LinkedIn.com
YouTube.com
Plus.Google.com
Tumblr.com

While this is not a complete list of active social networks, it represents the most popular social networks at the time of this writing. At this point, our custom search engine would search only these websites and provide all results integrated into one search result page. We now want to add refinements that will allow us to isolate the results from each social network.

After you have added these websites, navigate to the control panel option in order to view the configuration of this custom search engine. On the left menu, you should see an option called "Search Features". This will present a new option at the top of the page labeled "Refinements". Click the "add" button to add a new refinement for each of the websites in this example. You should create these in the same order that you want them to appear within the search results. For this demonstration, I created the following refinements in order.

Facebook
Twitter
Google+
Instagram
LinkedIn
YouTube
Tumblr

When each refinement is created, you will have two options of how the search will be refined. The option of "Give priority to the sites with this label" will place emphasis on matching rules, but will also reach outside of the rule if minimal results are present. The second option of "Search only the sites with this label" will force Google to remain within the search request and not disclose other sites. I recommend using the second option for each refinement.

Now that you have the refinements made, you must assign them each to a website. Back on the "Setup" menu option, select each social network website to open the configuration menu. Select the dropdown menu titled "Label" and select the appropriate refinement. Repeat this process for each website and save your progress. You should now have a custom search engine that will not only search several specific social network websites, but it should also allow you to isolate the results for each network. Navigate back to the control panel view and select the Public URL button to see the exact address of your new engine. Go to that address and you should see a very plain search engine. You can now search any term or terms that you want and receive results for only the social networks that you specified. Additionally, you can choose to view all of the results combined or only the results of a specific network. Figure 3.03 displays the results when I searched the term osint. In this example, I have selected the Twitter refinement in order to only display results from twitter.com.

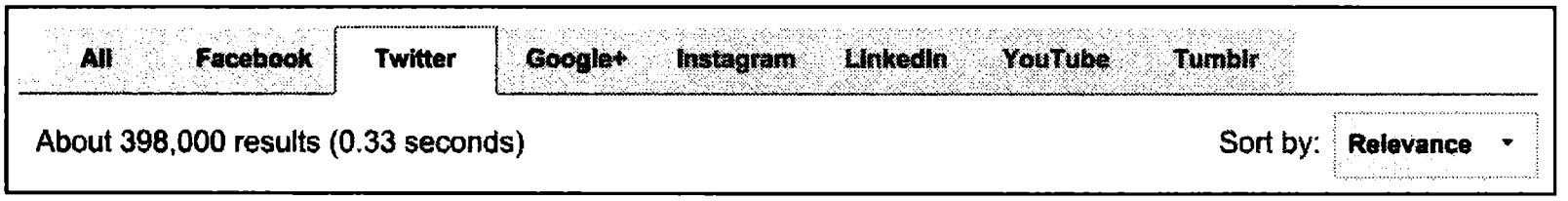


Figure 3.03: A Twitter refinement in a Google Custom Search.

You can now bookmark this new search engine that you created and visit it whenever you have a target to search. You can take your custom search engines to another level by adding refinements that are not website specific. In the next example, we will make a search engine that will search the entire internet and allow us to filter by file type.

Create a new custom search engine and title it "Documents". Add only "google.com" as the website to be searched. We do not actually want to search google.com, but a website is required to get to the control panel. Save your engine and open the control panel to configure the options. In the "Sites to search" portion, choose the "Search only included sites" option and change it to "Search the entire web but emphasize included sites" option. Delete google.com from the sites to be searched. You now basically have a custom search engine that will search everything. It will essentially do the same thing as Google's home page. You can now add refinements to filter your search results. Navigate to the search features menu and add a new refinement. Title the new refinement "PDF"; leave the default setting of "Give priority to the sites with this label"; and enter the following in the "Optional word(s)" field.

ext:pdf

This will create a refinement that will allow you to isolate only PDF documents within any search that you conduct. Save this setting and create a new refinement. Title it DOC; leave the default search setting; and place the following in the "Optional word(s)" field.

ext:doc OR ext:docx

This will create a new tab during your search results that will allow you to isolate Microsoft Word documents. By entering both the doc and docx formats, you will be sure to get older and newer documents. The word "OR" tells Google to search either format. Repeat this process for each of the following document types with the following language for each type.

XLS (Excel Spreadsheets) – ext:xls OR ext:xlsx OR ext:csv
PPT (PowerPoint Files) – ext:ppt OR ext:pptx
TXT (Text Docs) – ext:txt OR ext:rtf
WPD (Word Perfect Docs) – ext:wpd
ODT (Open Office Docs) – ext:odt OR ext:ods OR ext:odp
ZIP (Compressed Files) – ext:zip OR ext:rar OR ext:7z

Figure 3.04 displays the results of a search for the term osint within this new engine. The PPT tab is selected which reveals 45 PowerPoint presentations that contain the term. There are endless possibilities with this technique. You could make an engine that searched for audio and video files with extensions such as mp3, mp4, mpeg, avi, mkv, etc. You could make an engine that isolated images with extensions such as jpg, jpeg, png, bmp, gif, etc. You could also replicate all of this into a custom engine that only searched a specific website. If you were monitoring threats against your company, you could isolate only these files that appear on one or more of your company's domains.

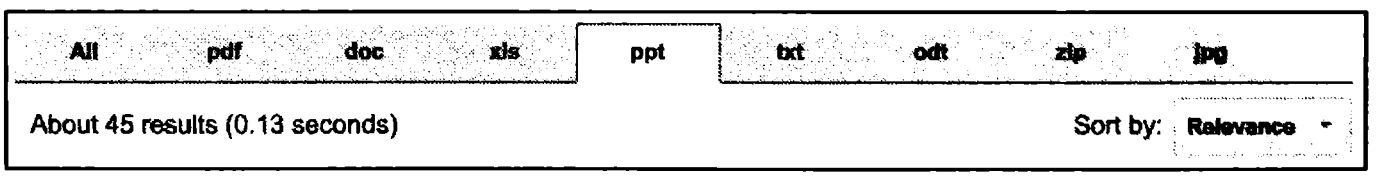


Figure 3.04: A PowerPoint file refinement within a Google Custom Search.

As a service to my readers, I have created several custom search engines that are publicly available. You can either connect to them through my website under the OSINT Links section, or navigate to the following addresses.

Social Networks: https://inteltechniques.com/OSINT/social.networks.html
Smaller Networks: https://inteltechniques.com/OSINT/smaller.networks.html
Dating Sites: https://inteltechniques.com/OSINT/dating.networks.html
Pastebins: https://inteltechniques.com/OSINT/pastebins.html
Cloud Documents: https://inteltechniques.com/OSINT/docs.html
Document Formats: https://inteltechniques.com/OSINT/docs.format.html

One negative aspect to custom Google search engines is that they only display the most relevant 100 results. These are presented in ten pages of ten results per page. If you are searching very specific terms, this may not be an issue. However, standard searches can be limiting. The fix for this is an amendment to the URL of the custom search engine. As an example, I will use the social network engine that was created earlier. The public URL of this engine is the following address.

https://www.google.com/cse/publicurl?cx=001580308195336108602:oyrkxatrfyq

Assume you had searched "OSINT Tools" within the Twitter option of this engine. Google would announce that 927 results are available. However, you can only view the first 100 within this custom engine. If you copy the search engine ID you can create a new address that will lift this limit to 1000 results. The following URL will expand this specific search engine.

www.google.com/custom?cx=001580308195336108602:oyrkxatrfyq&num=100&filter=0

While the look and feel is different, the results are the same. However, you can now view ten pages of 100 results each for a total of 1000 results.

Google Alerts (google.com/alerts)

When you have exhausted the search options on search engines looking for a target, you will want to know if new content is posted. Checking Google results every week on the same target to see if anything new is out there will get mundane. Utilizing Google Alerts will put Google to work on locating new information. While logged into any Google service, such as Gmail, create a new Google Alert and specify the search term, delivery options, and email address to which to send the alert. In one of my alerts, Google will send an email daily as it finds new websites that mention "Open Source Intelligence Techniques" anywhere in the site. Another one of my alerts is for my personal website. I now receive an email when another site mentions or links to my website. Parents can use this to be notified if their child is mentioned in a website or blog. Investigators that are continuously seeking information about a target will find this beneficial.

Talk Walker Alerts (talkwalker.com/en/alerts)

While Google Alerts are the current standard for email alerts on search terms, there are other options. Some people choose not to use Google because of privacy concerns. Some just want alternatives to obtain complete results. Talk Walker Alerts has a feel very similar to Google Alerts. The main difference is that it does not require a Google account. Applying both of these services to monitor keywords, email addresses, user names, or real names will generate the best results possible for your investigation.

Real World Application: A police detective was assigned a runaway case where a 15-year-old had decided to leave home and stay with friends at an unknown location. After several extensive

internet searches, a Google Alert was set up using the runaway's name and city of residence. Within three days, one of the alerts was for a blog identifying the runaway and where she was currently staying. Within 30 minutes, the unhappy child was back home.

Bing (bing.com)

Google is not the only great search engine. While Google is the overwhelming choice of search engines used today, other sites should not be ignored, especially when having trouble locating any information on a subject. Bing is Microsoft's competition to Google and provides a great search experience. In 2009, Yahoo search (yahoo.com) began using the Bing search engine to produce search results. This makes a Yahoo search redundant if a Bing search has already been conducted. The same tactics described previously, and in the numerous Google books, can be applied to any search engine. The "site" operator and the use of quotes both work with Bing exactly as they do with Google. Bing also introduced time filtered searching that will allow you to only show results from the last 24 hours, week, or month. There are a couple of additional operators that are important that only apply to Bing. Bing offers an option that will list every website to which a target website links, and is the only search engine that offers this service.

Bing LinkFromDomain

I conducted a search on Bing of "LinkFromDomain:inteltechniques.com". Note that there are no spaces in the entire search string and omit the quotation marks. This operator creates a result that includes every website that I have a link to on any of the pages within my website. This can be useful to an investigator. When a target's website is discovered, this site can be large and contain hundreds of pages, blog entries, etc. While clicking through all of these is possible, sometimes links are hidden and cannot be seen by visually looking at the pages. This operator allows Bing to quickly pull links out of the actual code of the website.

Bing Contains

Earlier, I discussed searching for files with specific file extensions on Google. The "filetype" and "ext" operators that were explained both work on Bing the same way. However, Bing offers one more option to the mix. The "contains" operator allows you to expand the parameters of the file type search. As an example, a Bing search of "filetype:ppt site:cisco.com" returns 3,200 results. These include PowerPoint files stored on the domain of cisco.com. However, these results do not necessarily include links on the cisco.com website to PowerPoint files stored on other websites. A search on Bing for "contains:ppt site:cisco.com" returns 7,200 results. These include PowerPoint files that are linked from pages on the domain of cisco.com, even if they are stored on other domains. This could include a page on cisco.com that links to a PowerPoint file on hp.com. In most cases, this search eliminates the need to conduct a filetype search, but both should be attempted.

Google Images (images.google.com)

Google Images scours the web for graphical images based on a search term. Google obtains these images based on keywords for the image. These keywords are taken from the filename of the image, the link text pointing to the image, and text adjacent to the image. This is never a complete listing of all images associated with a subject, and will almost always find images completely unrelated to a target. In the case of common names, one should enclose the name in quotes and follow it with the city where the subject resides, place of employment, home town, or personal interests. This will help filter the results to those more likely to be related to the subject. When results are displayed, clicking the "Search tools" button will present six new filter menus. This menu will allow you to filter results to only include images of a specific size, color, time range, or license type. The most beneficial feature of Google Images is the reverse image search option. This will be explained in great detail later in the book.

Bing Images (bing.com/images)

Similar to Google, Bing offers an excellent image search. Both sites autoload more images as you get toward the end of the current results. This eliminates the need to continue to load an additional page, and leads to faster browsing. Bing also offers the advanced options available on Google, and adds the ability to filter only files with a specified layout such as square or wide. Bing provides a "filter" option in the far right of results that provides extended functionality. The People tab offers restriction for images of "Just faces" and "Head & shoulders". It also provides suggested filters with every image search. Clicking image search links may provide additional photographs of the specific target based on the listed criteria. This intelligence can lead to additional searches of previously unknown affiliations.

Web Archives

Occasionally, you will try to access a site and the information you are looking for is no longer there. Maybe something was removed, amended, or maybe the whole page was permanently removed. Web archives, or "caches" can remedy this. I believe that these historical copies of websites are one of the most vital resources when conducting any type of online research. This section will explain the current options in order from most effective to least.

Google Cache (google.com)

When conducting a Google search, notice the result address directly below the link to the website. You will see a green down arrow that will present a menu when clicked. This menu will include a link titled "Cached". Clicking it will load a version of the page of interest from a previous date. Figure 3.05 (first image) displays a search for phonelosers.org which returns a result that includes a cached version of the page. This version was taken four days prior to the current date, and displays information different from the current version. The second option visible within this menu, titled Similar, identifies web pages that contain content similar to the listed result.

If you have a specific page within a website that you want to view as a cached version, type the exact website into Google to link to the cached page. For example, if I wanted to see a previous view of the podcast for The Phone Show, an audio archive about telephone pranks, I would conduct a Google search for the site "www.phonelosers.org/snowplowshow". This will return the main landing page as well as sub-pages that will each have a cached view. If any of these pages were to go offline completely, Google would hold the last obtained version for viewing. I could have also typed the following directly into any Google search page to be navigated directly to the cached page.

cache:www.phonelosers.org/snowplowshow

Bing Cache (bing.com)

Similar to Google, Bing offers a cached view of many websites. Searching for a domain name, such as phonelosers.org will present many results. The first result should link to the actual website. Directly next to the website name is a small green down arrow. Clicking it will present the option of "Cached page". Clicking this link will display a previous version of the target website as collected by Bing. Figure 3.05 (second image) displays their menu option.

Yandex Cache (yandex.com)

The Russian search engine Yandex will be explained in great detail later, but it is important to note now that it also possesses a cache option. Very similar to Google and Bing, Yandex presents a green drop-down menu directly under the title of the search result. Figure 3.05 (third image) displays their cache menu option. Selecting the Cached page option opens a new tab displaying the most recent Yandex archive of the page. The top banner displays the date and time of capture, the original website address, and a search option to highlight selected keywords within the result. The biggest strength of the Yandex cache is the lack of updates. While this may sound counterintuitive, an older cache can be very helpful in an investigation. Assume that the Phone Losers website was your target. At the time of this demonstration, January 10, 2016, the Google, Bing, and Yandex caches of this page were dated as follows.

Google: January 10, 2016
Bing: January 9, 2016
Yandex: January 2, 2016

Google and Bing tend to have very recent results which often appear identical to the live view. However, the Yandex option from a week prior is more likely to contain modified content. You can often locate a cached version of a page that is older than the Yandex version on Baidu.

Baidu Cache (baidu.com)

This Chinese search engine is the least productive as far as cached copies of websites, but it

should not be ignored. It will be explained further during a later discussion about international engines. The results of a search on Baidu are mostly in Chinese, but can still be valuable to those that cannot read the text. At the bottom of each search result is a green link to the website that hosts the content of the result. While this also includes a drop-down menu, the cache option is not there. Instead, look for a word in Chinese directly to the right of this link. In Figure 3.05 (fourth image) it is displayed as 百度快照. Clicking this link will open a new tab with the cache result, which Baidu refers to as a snapshot. In my experience, the presence of this linked option does not always mean that a cached version exists.

Archive.is (archive.is)

There are two additional options for historic archives of web pages. However, neither of them is very powerful and my successes with these services have been minimal. Archive.is will allow you to search a domain and display any captured archives of the home page. They also offer a wildcard service that will search an entire domain for any captured pages. If I wanted to search inteltechniques.com for any archived pages, I would enter the following two queries into the search field.

https://inteltechniques.com/*
*.inteltechniques.com

The first would look for any archived pages within the chosen domain. The second would look for any sub-domains such as mail.inteltechniques.com or ftp.inteltechniques.com. I have rarely received results from the sub-domain search through this service. The strength of Archive.is is their stance on ignoring the noarchive rule of the robots.txt file. This file will be fully explained in a later chapter. In brief summary, it is a set of rules included on a web server that instructs search engines on allowable ways to index the website. The noarchive function tells a search engine to never archive anything on the site. As an example, my website inteltechniques.com has this enabled. Therefore, Google, Bing, Yandex, and Baidu do not have a cache of any pages on that domain. However, Archive.is does, and they have multiple versions from the past three years. Consequently, this service should always be checked when dealing with a tech savvy target that blocks traditional engines from caching a page.

Coral (coralcdn.org)

This service offers a different spin on web archiving. Instead of a collection of historic caches of a website, it retrieves snapshots of any website that might not be obtainable by your connection. I have used this many times in the past when I cannot connect to a website due to overload on the providing server. This could happen when a small website receives an abundance of visits due to instant popularity on a service such as Reddit. If I cannot load the desired page, I will attempt to view it through Coral. To do this, I append the website address with ".nyud.net". If I were attempting to view inteltechniques.com, I would instead navigate to the exact address of inteltechniques.com.nyud.net. This will generate a current snapshot of the target page. It is

obtained by a series of servers that have the power to retrieve online content that is otherwise restricted due to high demand.

The Wayback Machine (archive.org/web/web.php)

The Wayback Machine will provide a much more extensive list of options in viewing a website historically. Searching for phonelosers.org displayed a total of 1,197 captures of the site dating from 12/21/1997 through 12/28/2017 (Figure 3.06). Clicking the links presents quite a display of how the site has changed. Graphics are archived as well, proving that we should always think twice about which photos we post to the internet. Each view of the archived page will allow the user to click through the links as if it were a live page on the original web server. Clicking through the timeline at the top of each page will load the viewed page as it appeared on the date selected.

Wayback Search

Until 2016, you could not search keyword across Wayback Machine data. You had to know the exact URL of a target website, or at least the domain name. Today, we can search any terms desired and connect directly to archived data. At the time of this writing, a search bar was present at the top of every Wayback Machine page. If that should change, you can also conduct a search via a direct URL. The following address searched "Michael Bazzell" throughout the entire archive of information.

https://web.archive.org/web/*/Michael Bazzell

The results identify over twenty websites that include these terms. Within those sites are dozens of archived copies of each. This data represents decades of content at your fingertips. Much of it is offline and unavailable on the current public internet. Many domains have completely shut down. Furthermore, websites that I own appear within the results, even though I have specifically blocked archiving them through a configuration file on my server. You would not find these by searching the domains directly through the Wayback Machine. This is a reminder that we should check all available resources before completing our investigations.

Searching All Resources

There are occasionally websites that surface claiming to be able to extract and rebuild entire websites from online caches. In my experience, none of these have ever provided a complete historical view versus a manual approach. Engines such as Bing and Yandex generate a unique code when a cache is displayed. This action prevents most automated search tools from collecting archived information. I do not believe any option, other than navigating to each resource, will present you with the content that you need. I bookmark each of these services in an individual folder titled Archives and open each tab when I have a domain as a target. I have also created an online tool that will collect your target domain and forward you to the appropriate archive page. This will be explained later in Chapter Sixteen when discussing domain searches.

Finally, it is important to acknowledge that these resources can be beneficial when everything on a website appears to be present and unaltered. While caches work well on websites that have been removed and are completely empty, they also can tell a different story about websites that appear normal. Any time that I find a website, profile, or blog of interest, I immediately look at caches hoping to identify changes in content. These minor alterations can be very important. They highlight information that was meant to be deleted forever. These details can be the vital piece of your investigation puzzle. Most people have no idea that this technique exists.

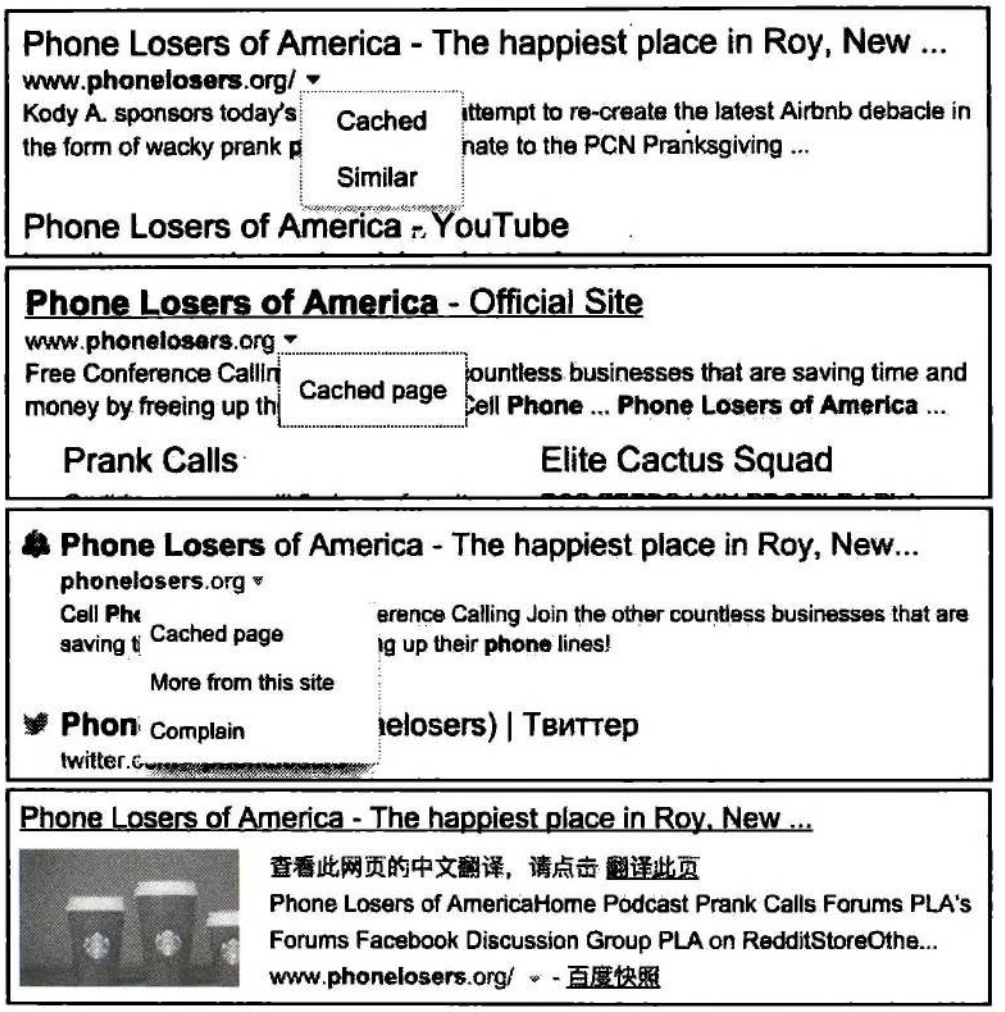


Figure 3.05: Cache menu options on Google, Bing, Yandex, and Baidu.

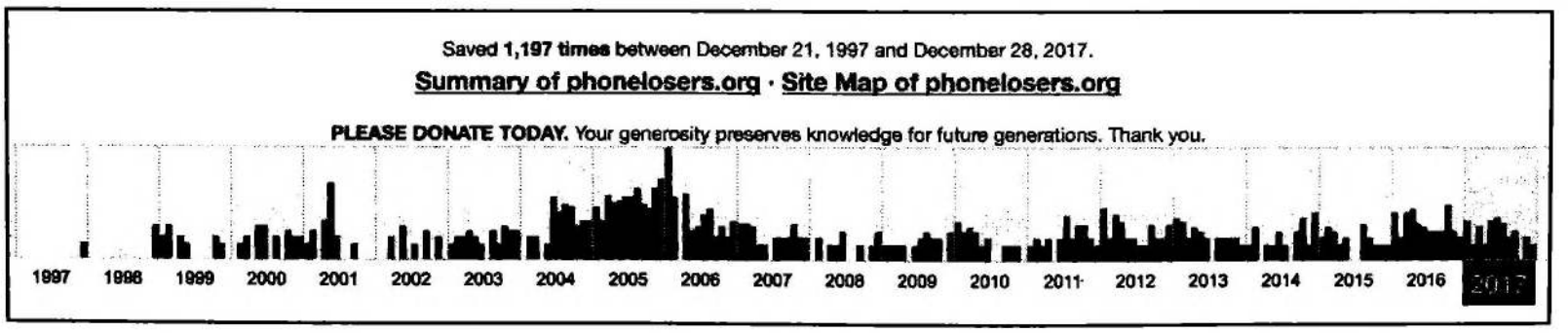


Figure 3.06: Wayback Machine results for an archived website.

Google Translator (translate.google.com)

Many websites exist in non-English languages. As internet enthusiasts, we tend to focus on sites within our home area. There is a wealth of information out there on sites hosted in other countries which are presented in other languages. Google Translator will take text from any site or document and translate the text to a variety of languages. Usually, the service will automatically identify the language of the copied and pasted text. Selecting the desired output will provide the translation.

Alternatively, you can translate an entire website in one click which will give a native view of the layout of the site. Instead of copying individual text to the search box, type or paste in the exact URL (address) of the website you want translated. Clicking the "Translate" button will load a new page of the site, which will be translated to English. This translation is rarely, if ever, perfect. However, it should give you an idea of the content presented on the page. This will also work on social network sites such as Twitter and Instagram.

Bing Translator (bing.com/translator)

A few years after Google introduced free translation services, Bing created their own product. At first glance, it looks like a replica of Google's offering. However, Bing's translations are usually slightly different than Google's results. Similar to Google, you can also type or paste an entire foreign website to conduct a translation of everything on the target page. This is where the similarities stop. After you translated an entire page and see the results in English, you can control the view and include both the original and translated text. The default view will display the original text as you hover your cursor over any translated text. Another option allows you to view the page as it originally appeared and hover over text to see the translation. Additional options allow you to view a top and bottom or side by side translation. These view options can be extremely beneficial when documenting evidence. It may help other viewers that are not bilingual digest the content more easily. When I am using a high-resolution external monitor, I always chose the "Side by side" option as seen in Figure 3.07.

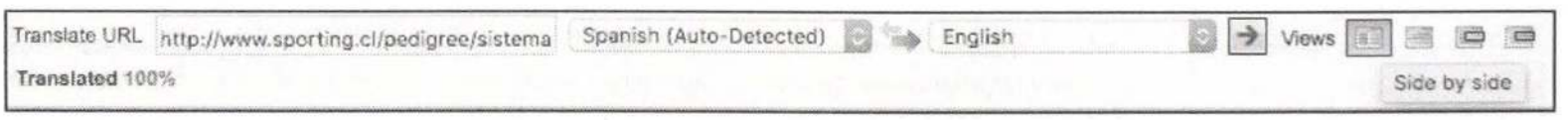


Figure 3.07: Translation options within a Bing Translator result.

Online Translator (online-translator.com)

There are dozens of additional online translation tools available. Almost all of them allow translation of a small amount of text at a time. Some use either the Google or Bing translation service. One last online translation tool worth mentioning is Online Translator. It is unique from the dozens of other options in that it allows translation of entire websites similar to Google and Bing. This service provides an independent translation and can be considered a third source.

I am often asked during training which of the three services I use during investigations. My answer is all three. This is important for two reasons. The obvious benefit is that you will receive three unique translations that will be very similar. The minor variations may be important, especially when translating Tweets and other shortened messages that may not be grammatically correct in any language. The second reason is to show due diligence during my investigation. I always want to go above and beyond what is required. Translating a foreign web page through three different services emphasizes my desire to conduct an unbiased investigation.

Non-English Google Results

Not every piece of information that will be useful to you will be obtained by standard searches within English websites. Your target may either be from another country or have associates and affiliations in another country. While Google and Bing try to pick up on this, the technology is not perfect. Google has a search site and algorithm that change by location. For example, google.fr presents the French search page for Google. While this may produce the same overall results, they are usually in a different order than on google.com. Google maintains a page with links to each international version of its search at google.com/language_tools. This can allow you to search each site for variations, but I have a preferred method.

2Lingual (2lingual.com)

This page will allow you to conduct one search across two country sites on Google. The Google search will display a plain search box and choices of two countries. The results will display in single columns next to each other. Additionally, the foreign results will be automatically translated to English. This feature can be disabled, if desired. The first few sponsored results (ads) will be similar, but the official results following should differ. This site can also be helpful when demonstrating to someone the importance of searching targets through multiple countries.

Google Input Tools (google.com/inputtools/try)

There is one last feature regarding foreign language searching that I have found useful. Google's Input Tools allow you to type in any language you choose. Upon navigating to the above website, choose the language of your target search. In Figure 3.08, I have chosen Arabic as the language and typed "Online Investigation" on a standard English keyboard. The result is how that text might appear in traditional Arabic letters. I have had the most success with this technique on Twitter. When supplying any search term on Twitter, the results are filtered by the presence of the keywords entered and only in the language provided. Searching "Online Investigation" on Twitter only provides results that have that exact spelling in English characters. However, searching the Arabic output provides Tweets that include the Arabic spelling of the selected words. This technique is extremely important when you have located a user name in a foreign language. As with all computer-generated translation services, the results are never absolutely accurate. I expect this technology to continue to improve.

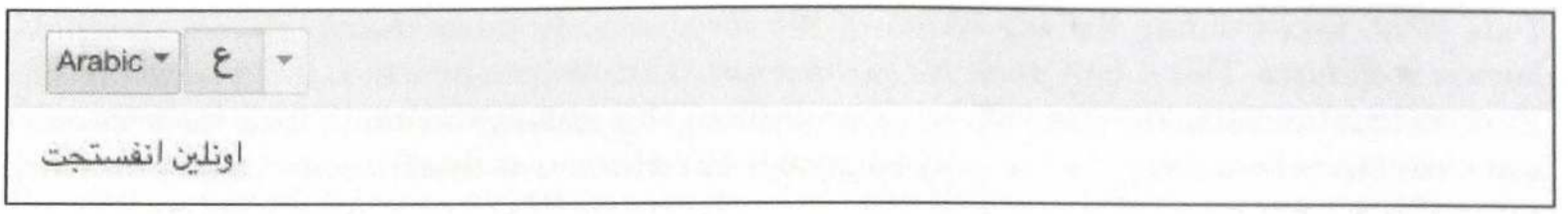


Figure 3.08: A Google Input Tools translation from English to Arabic.

Google Groups (groups.google.com)

Google Groups provides access to both Usenet groups and Non-Usenet Google groups. Usenet groups are similar to mailing lists. The Usenet archive is complete and dates back to 1981. Since many people posted to these groups using their real name or email address, identifying their opinions on controversial topics is effortless. Additionally, searching a real name will often provide previous email addresses that may not be known to the searcher. This provides new intelligence for future searches. While none of this is usually damaging to the submitter, it helps provide an overall view of the target of interest. Many of the newer groups used are created through Google and conform to practically any interest imaginable. Most users continue to use a real name, screen name, email address, or a combination of all three. Searching these posts is similar to any Google search. I suggest leaving the "all groups" option checked unless you are only looking for a specific Google Group.

Yahoo Groups (groups.yahoo.com)

While Google Groups will search many of the non-Google online groups, it will not pick up all of the Yahoo Groups. The content of most Yahoo Groups is public and will allow viewing without membership. Searching by real name or screen name will often produce results. Any time I have a target with a Yahoo email address, I search the user name before @yahoo.com through Yahoo Groups. Many people that possess a Yahoo email address have had it for many years. Yahoo Groups were very popular in the early days of the internet. In my experience, you are more likely to find content from Yahoo users in Yahoo Groups than Gmail users in Google Groups.

I have had multiple successes with searches in Google and Yahoo Groups. Most have been associated with pedophiles or background checks. In pedophile cases, I have identified new evidence based on historic conversations in various Yahoo Groups. This never identified new victims or generated new cases, but the details strengthened the current charges by showing a pattern of inappropriate interest in children. With background checks, this content has been extremely valuable. While applicants can easily clean their blogs and social profiles, they cannot easily purge their history within these groups. Many people simply forgot about the content, which was often posted a decade earlier. In one example, an applicant had signed an affidavit claiming to have never abused any type of narcotics. However, a post to a newsgroup indexed by Google revealed that he had battled cocaine addiction for several years.

Google News Archive (news.google.com)

This can be an amazing resource of information about a target. In the past, if someone relocated to a new geographical area, he or she could leave the past behind and start over. Today, that is difficult. Google's News Archive is continually adding content from both online archives and digitized content from their News Archive Partner Program. Sources include newspapers from large cities, small towns, and anything in between. The link referenced above will allow for a detailed search of a target's name with filters including dates, language, and specific publication. In order to display this menu, click on the down arrow to the right of the search box. This can quickly identify some history of a target such as previous living locations, family members through obituaries, and associates through events, awards, or organizations.

Google Newspaper Archive (news.google.com/newspapers)

The previous option focused solely on digital content, such as your local newspaper website. Google's Newspaper archive possesses content from printed newspapers. All results on this site consist of high-resolution scanned newspaper pages. In my experience, this collection is not as extensive as the next option discussed. However, it is definitely worth a look, and will likely continue to grow.

Newspaper Archive (newspaperarchive.com)

This paid service provides the world's largest collection of newspaper archives. The high resolution PDF scans of entire daily newspapers range in date from the 1800's until present. The first four editions of this book explained a method of using the Google Site operator and cached results to obtain practically any page of this newspaper collection without paying or subscribing. These vulnerabilities have all been patched and none of those techniques work today. Fortunately, Newspaper Archive still offers a 14-day free trial with unlimited access to every archive. While multiple trials can be obtained, each require a unique credit card number and email address. Many libraries have asked this service to scan their entire microfilm archives and make them freely available online. You will not find any mention of this free alternative on their home page, but a bit of searching will guide you to the right place. The following search on Google identifies hundreds of public libraries that pay for your access to their archives.

site:newspaperarchive.com "This archive is hosted by" "create free account"

The first part of the search tells Google to only look at the website newspaperarchive.com. The second part mandates that the exact phrase "This archive is hosted by" appears in the result. The final piece isolates only the newspaper collections that are available for free, and without a credit card. This identifies the landing pages of the various libraries that have made their collections freely available. While you will still be required to register through the service, payment is not required for these collections. Consider the following usage that will likely present you with free views of Newspaper Archive whenever you need them.

On 12/13/2017, I navigated to newspaperarchive.com/advancedsearch/ and conducted an advanced search for anyone named Michael Williams from Cedar Rapids, Iowa. Newspaper Archive presented several results from the Cedar Rapids Gazette. Clicking on any of these results prompted me to create an account and forced me to enter a valid credit card number to proceed. I could not create an account from any of the pages without providing payment. Instead, I conducted the following Google search.

site:newspaperarchive.com "This archive is hosted by" "cedar rapids gazette"

The first result was a direct connection to crpubliclibrary.newspaperarchive.com. Clicking this link presented a page dedicated to searching over 40 newspapers within the Cedar Rapids and Des Moines areas. In the upper right corner was a link titled "Create Free Account". I clicked this link and provided generic details and a throwaway email address. The membership choices now include a completely free option, which will only allow access to the Iowa newspapers. After creating my free account, I returned to crpubliclibrary.newspaperarchive.com and repeated my search of my target. Every link allowed me full unrestricted access to the high-resolution images.

While still logged into this account, I navigated to delawarecolib.newspaperarchive.com, the direct page associated with the Delaware County Library (which I found through the original Google search in this section). I was not authorized to view this newspaper collection. However, after clicking "Create Free Account" on this page, I entered the same data as provided to the Iowa newspaper previously. After verifying my email address, I was allowed immediate access to this series of newspapers.

This technique will not obtain access to every collection on Newspaper Archive. However, it will provide a surprising amount of free access to huge collections internationally. During an hour of downtime, I created a free account on every library collection I could locate, using the same credentials on each. I can now log into my single Newspaper Archive account and navigate the site from any page. When I reach a newspaper of interest after a search, I will be given full access if it is within a free collection. This is all thanks to the local libraries that have paid this site to give free access to the public. If the free trial of Newspaper Archive or the free library collections do not offer enough content, consider the following options.

Old Fulton (fultonhistory.com/Fulton.html): 34,000,000 scanned newspapers from the United States and Canada.

Library of Congress US News Directory (http://chroniclingamerica.loc.gov/):
Scanned newspapers from the United States dated 1836-1922.

Library of Congress US News Directory (http://chroniclingamerica.loc.gov/search/titles/):
Scanned newspapers from the United States dated 1690-Present.

Google Advanced Search (google.com/advanced_search)

If the search operators discussed in this chapter seem too technical, Google offers an advanced search page that simplifies the process. Navigating to the above website will present the same options in a web page that are possible by typing out the operators. This will help you get familiar with the options, but it will be beneficial to understand the operators for later use. The Advanced Search page will allow you to specify a specific phrase for which you are searching, just like the quotes in a search will allow. The site and filetype operators used earlier can be achieved by entering the desired filters on this page. It should be noted that the file type option on this page is limited to popular file types, where the filetype operator can handle many other file extensions.

Bing Advanced Search (search.yahoo.com/web/advanced)

Bing does not technically provide an advanced search page similar to Google's. However, since Yahoo uses Bing's search, you can use Yahoo's advanced search page as a replacement. This page will allow you to easily create a search that filters by individual terms, exact phrases, omitted terms, specific domains, file formats, and languages.

Additional Google Engines

Google isolates some search results into specialized smaller search engines. Each of these focuses on a unique type of internet search. The following engines will likely give you results that you will not find during a standard Google or Bing search. While some results from these unique searches will appear within standard Google results, the majority will be hidden from the main page.

Google Blogs (google.com)

Google removed its original blog search in 2014. It was quite helpful and focused mostly on personal websites, especially those with a blogging platform. Today, this is nothing more than a subsection of Google News. You can load the Blogs option under the "All News" menu within the Tools option on any Google News results page. Alternatively, you can navigate to the following address, replacing TEST with your search terms.

google.com/search?q=TEST&tbm=nws&tbs=nrt:b

The website above displays a standard Google search option, but the results appear much differently. A standard Google search of my name reveals my website, Twitter, and Amazon pages in the first results. The Google Blogs option reveals several personal and professional (media) blogs that mention my name. These results are likely buried within the standard Google search.

Google Patents (google.com/webhp?tbm=pts)

Google probably has the best patent search option on the internet. It allows you to search the entire patent database within any field of a patent. This can be useful for searching names associated with patents or any details within the patent itself. If you need further help, Google offers an advanced patent search at google.com/advanced_patent_search.

Google Scholar (scholar.google.com)

Google Scholar is a freely accessible web search engine that indexes the full text of scholarly literature across an array of publishing formats. It includes most peer-reviewed online journals of Europe's and America's largest scholarly publishers, plus many books and other non-peer reviewed journals. My favorite feature of this utility is the case law and court records search. I have located many court records through this free website that would have cost money to obtain from private services.

Advangle (advangle.com)

This is a simple and convenient builder of complex web search queries for both Google and Bing. The service allows you to quickly build a query with multiple parameters, such as the domain, language, or date published, and immediately see the result of this query in Google or Bing search engines. You can save your queries in an Advangle account if you want to restore a search to identify new results. Any condition in a query can be temporarily disabled without removing it completely. This allows you to quickly try several combinations of different conditions and choose the most suitable for your needs. Figure 3.09 displays the search page with filters for my exact name, on my website, within the past month, and only PDF files. The Google and Bing "Open" options will launch a new tab with the exact terms required for these options.

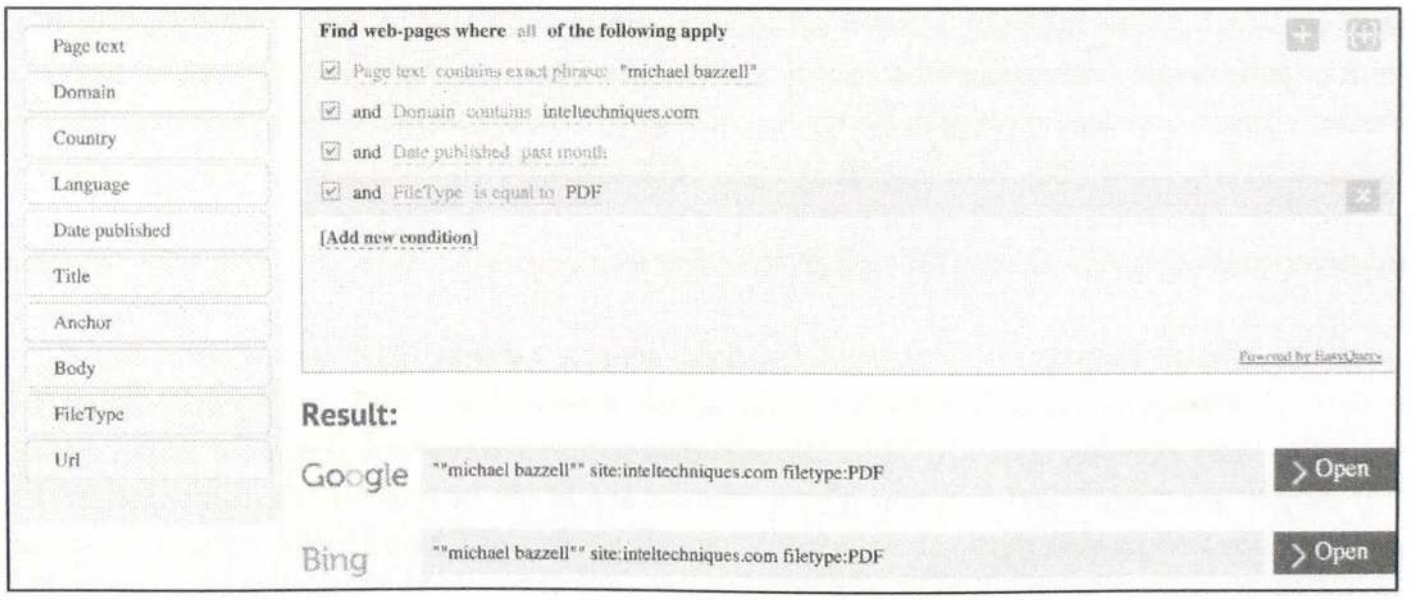


Figure 3.09: An Advangle search menu in use.

Keyword Tool (keywordtool.io)

Keyword Tool displays autocomplete data from Google, Bing, YouTube, and the App Store. You have likely noticed that Google quickly offers suggestions as you type in your search. This is called autocomplete. If I were to type "macb" into Google, it would prompt me to choose from the most popular searches when people typed those letters. This information may lead you to new terms to search in reference to your investigation. The advantage of Keyword Tool over Google is that Google only provides the ten most popular entries. Keyword Tool provides the ten most popular entries. Additionally, you can choose different countries to isolate popular terms. You can also see results from similar searches that Google does not display.

Real World Application: I have successfully used this technique during the investigation of many businesses. I was once asked by a medium sized business to investigate reports of a faulty product that they had recently recalled. They wanted to see customer complaints. After searching the typical review websites, I conducted a search with Keyword Tool. I discovered that the 9th most popular search involving this specific product name included a term that was a misspelling of the product name. It was different enough in spelling that my searches were missing this content. Knowing this information, I was able to locate more relevant data for the client.

Other Alternatives

Google and Bing are great, but they do not do it all. There will always be a need for specialized search engines. These engines usually excel in one particular search method which justifies the lack of search power in other areas. The sites listed in this next section represent the extreme minority when it comes to search traffic. It is often sites like these that implement the technologies that we later take for granted in more popular engines.

iSEEK (iseek.com)

One site that can be of assistance is iSEEK. This site provides categories of search results based on the information stored about the target. A search for "Glenn McElhose" provides search results similar to Google and Bing; however, there is an additional feature. The left column of the screen includes categories created by the search results. The results include topics, people, places, and organizations related to the target. Clicking these categories will filter the search results to only match the topic, person, place or organization that is selected. Clicking the selection again will un-filter the results. This can be very helpful when the target has a common name. The results can be overwhelming, but the categories may provide a filter that can be applied that will make the results more manageable.

Carrot2 (carrot2.org)

This is another search engine that groups search results into sets of topics. Similar to iSeek, it displays clustered results which allow you to navigate through large numbers of search results

quickly. One way that this service stands out from the others is that it offers three different layout formats for viewing and filtering the results. The initial search will produce categories in a left menu similar to the previously mentioned sites. The "Circles" and "Foam Tree" tabs will change these text options into interactive graphics. The size of each section correlates to the amount of search results that fit into that topic. In one example, a search of my name reveals the largest topic relating to me as "OSINT Training". Clicking that piece filters the search results to only include related entries to the right.

Exalead (exalead.com)

Headquartered in Paris, this search engine has gained a lot of popularity in the United States. The main search engine provides many results on popular searches. I have found that individual targets without a strong internet presence do not get many, if any, results on this site. However, this site excels in two areas. It works well in finding documents that include the target mentioned within the document. The "filetype" operator used in other engines works the same here. The document search option was also included in FOCA, the metadata scraper that will be discussed later in this book. Voxalead, an Exalead search engine, searches within audio and video files for specific words. This is thanks to speech to text technologies. Voxalead will search within all of the spoken audio of a file for references to the text searched. The results are presented in a timeline view. Currently, the majority of the results of this new product link to news media and public news video files.

Searx (searx.me)

This is considered a meta-crawler, as it presents results from both Google and Bing. It often gets dismissed as another comparison search site, but there are many other advantages to using this service. First, conducting a search will provide results from the two main search engines, but will remove duplicate entries. This alone is a quick way to conduct your due-diligence by checking Google and Bing. Next, the top row of options will allow you to repeat this redundancy-reducing option by checking results on Google's and Bing's Images, News, and Videos sections. Next to each result on any search page is a "cached" link. Instead of opening the Google or Bing cache, clicking this will open the cached page of the target website through the Wayback Machine. Finally, a "proxied" option next to each result will connect you to the target website through a proxy service provided by Searx. This is basically a layer of privacy preventing the website owner from collecting data about you, such as your IP address. Technically, Searx.me opened the target site, and their data would be tracked instead of yours. There are ways for adversaries to bypass this "anonymity", but it is decent protection for most sites.

The final benefit of this service over all others is the easy ability to export search results as a file. The "Links" section to the right of all search pages displays options to download a csv, json, or rss file of the results. The csv option is a simple spreadsheet that possesses all of the search results with descriptions and direct links. I find this helpful when I have many searches to conduct in a short amount of time, and I do not have the ability to analyze the results until later.

Million Short (millionshort.com) & **Million Tall** (milliontall.com)

These websites offer a unique function that is not found on any other search engine. You can choose to remove results that link to the most or least popular one million websites. This will eliminate popular (or unpopular) results and focus on lesser known websites (or well known, if desired). You can select to remove (or include) the top 100,000, 10,000, 1,000, or 100 results.

Tor Search Engines

Tor is free software for enabling anonymous communication. The name is an acronym derived from the original software project name The Onion Router. Tor directs internet traffic through a free, worldwide, volunteer network consisting of more than six thousand relays to conceal a user's location and usage from anyone conducting network surveillance or traffic analysis. Using Tor makes it more difficult for internet activity to be traced back to the user. This also applies to a website that is hosted on the Tor network. Usually, these sites include illegal drug shops, child pornography swaps, and weapon sales. Because these sites are not hosted on publicly viewable networks, they are hard to locate and connect. Two search engines and a proxy aid this process.

Ahmia (ahmia.fi)

This is a very powerful Tor search engine. While no engine can index and locate every Tor website, this is the most thorough option that I have seen. It should be the first engine used when searching Tor related sites. The links in the results will not load if searching through a standard browser and connection. Using the Tor Browser discussed in Chapter Two is the ideal way to use this service.

Onion Link (onion.link)

Similar to Ahmia, Onion Link attempts to identify websites within the Tor network. It uses a Google Custom Search Engine (CSE) and appends ".link" to each search result. This allows you to open these links through Onion Link's own Tor connection which appears to pull the data from their own cached sources. This makes the viewing of these pages much quicker with faster page loads, and eliminates the need to be on the Tor Browser. While relying on Google to index these pages is a bit amateur, the minimal results here are often different than other Tor engines. An alternative option which appears to index the same pages can be found at torchtorsearch.com.

Tor2Web (tor2web.org)

Whenever you see a URL like libertygb2nyeyay.onion, it is a Tor Onion website. As mentioned earlier, you cannot connect directly to these links without being connected to the Tor network. However, you can replace ".onion" within the address to ".onion.to" in order to view the content. In the above example, navigating to the website libertygb2nyeyay.onion.to will display the content using the Tor2Web proxies. This connects you with Tor2web, which then talks to the onion

service via Tor, and relays the response back. This is helpful when locating Tor links on Twitter.

Tor Scan (torscan.io)

This engine does not query terms and attempt to find matching Tor sites. Instead, it queries an identified Tor site and displays various page details. Searching the Tor URL from the previous example displays the first date the address was seen, the last time it was scanned, any IP addresses associated with it, the index page source code, and the metadata from the file. I now know that our target site is hosted on 172.19.0.13, was last modified on 08/17/2017, and was present earlier today. This query was conducted without the need to connect to the Tor service. This service is still in "Beta", but shows great promise. It could potentially become the Wayback Machine for Tor. I now use this site as my first stop when I have a target Tor (.onion) address.

Tor Search Sites

I believe some of the strongest Tor search engines exist only on the Tor network. You cannot access them from a standard internet connection, and the Tor Browser is required for native use. My favorite is "Not Evil", which can be found at the following address if connected to Tor.

hss3uro2hsxfogfq.onion

Since Tor2Web allows us to use their proxy, we can connect to "Not Evil" by navigating directly to the following Tor2Web proxy address, without being on the Tor Browser.

hss3uro2hsxfogfq.onion.to

This presents the home page of the search site, and allows for a keyword search. However, searching through this portal while being connected through the Tor2Web proxy can present difficulties. Instead, consider conducting a search within a URL submission. In the following web address, I am connecting to Tor2Web's proxy of the search engine and requesting a search results page for the term OSINT.

hss3uro2hsxfogfq.onion.to/index.php?q=OSINT

This type of submission will be much more reliable than counting on the proxy to conduct your search and return an additional proxy-delivered page.

International Search Engines

Search engines based in the U.S. are not the primary search sites for all countries. Visiting search sites outside of the U.S. can provide results that will not appear on Google or Bing. In Russia, Yandex is the chosen search engine. Yandex offers an English version at yandex.com. These results are often similar to Google's; however, they are usually prioritized differently. In the past,

I have found unique intelligence from this site when Google let me down. In China, most people use Baidu. It does not offer an English version; however, the site is still useable. Striking the "enter" key on the keyboard after typing a search will conduct the search without the ability to understand the Chinese text. New results not visible on Google or Bing may be rare, but an occasional look on these sites is warranted.

Yandex (yandex.com)

In a previous edition of this book, I only made a brief reference to Yandex and quickly moved on. In the past two years, I had discovered many advanced features of Yandex which justify an expanded section. Visually, the Yandex home page and search result pages do not possess additional search operators. These options are only available by issuing a direct command within your search. While this can be more cumbersome than a Google search, the results can include much new data. Some of these searches can be overkill for daily use, but those who conduct brand reputation monitoring or extensive background checks may take advantage of this.

Exact terms: Similar to Google and Bing, quotation marks will search for exact terms. Searching "Michael Bazzell" inside of quotes would search those terms, and would avoid "Mike" or "Bazel".

Missing word: You can search an exact phrase without knowing every word of the phrase. A search for "Open Source * Techniques" inside of quotation marks will identify any results that include that phrase with any word where the asterisk (*) is located. This identified not only results with the title of this book, but also results for "Open Source Development Techniques" and "Open Source Responsive Techniques". This search can be very useful for identifying a person's middle name. "Michael * Bazzell" produced some interesting results.

Words within the same sentence: The ampersand (&) is used in this query to indicate that you want to search for multiple terms. "Hedgehog & Flamingo", without the quotation marks, would identify any websites that contained both of those words within one sentence. If you want the results to only include sentences that have the two words near each other, you can search "Hedgehog /2 Flamingo". This will identify websites that have a sentence that includes the words Hedgehog and Flamingo within two words of each other.

Words within the same website: Similar to the previous method, this search identifies the searched terms within an entire website. "Hedgehog && Flamingo", without quotation marks, would identify pages that have both of those words within the same page, but not necessarily the same sentence. You can also control the search to only include results that have those two words within a set number of sentences from each other. A search of "Hedgehog && /3 Flamingo", without the quotation marks, would identify websites that have those two words within three sentences of each other.

Include a specific word: In Google and Bing, you would place quotation marks around a word to identify pages that contain that word in them. In Yandex, this is gained with a plus sign (+).

Michael +Bazzell would mandate that the page has the word Bazzell, but not necessarily Michael.

Search any word: In Google and Bing, you can use "OR" within a search to obtain results on any of the terms searched. In Yandex, this is achieved with the pipe symbol (|). This is found above the backslash (\) on your keyboard. A search of "+Bazzell Michael|Mike|M ", without quotation marks, would return results for Michael Bazzell, Mike Bazzell, and M Bazzell.

Exclude a word: Google and Bing allow you to use a hyphen (-) to exclude a word in a search. Yandex does not technically support this, but it seems to work fine. The official Yandex operator is the tilde (~). A typical search would look like "Michael Bazzell ~ Mike", without the quotation marks. This would identify websites that contained Michael Bazzell, but not Mike Bazzell. I prefer to stick with the hyphen (-) until it no longer works.

Multiple identical words: This is a technique that I have needed several times in the past before I learned of Yandex's options. You may want to search for websites that contain a specific word more than once. An example might be if you are searching for someone that has two identical words in his or her full name. "Carina Abad Abad" would fit in this scenario. You could use quotation marks to identify the majority of the results, but you would filter out anything that was not exact such as Abad,Abad, Abad-Abad, or AbadAbad. This is where the exclamation point (!) comes in. A search of "!Carina !Abad !Abad", without quotation marks, would identify any results that included those three words regardless of spacing or punctuation.

Date specific searches: While Google provides a menu to filter your searches by date, Yandex makes you work harder for it. You must specify the date range within the search. The following queries should explain the options.

date:20111201..20111231 OSINT – Websites mentioning OSINT between December 1-31, 2011
date:2011* OSINT – Websites mentioning OSINT in the year 2011
date:201112* OSINT – Websites mentioning OSINT in December of 2011
date:>20111201 OSINT – Websites mentioning OSINT after December 1, 2011

Standard operators: Most of the operators explained earlier for Google and Bing should also work in Yandex. The commands for Site, Domain, Inurl, and Intitle should work the same way. Yandex maintains a list of operators at https://yandex.com/support/search/how-to-search/search-operators.html. All Yandex operators work together and multiple operators can be used to form very specific searches. Figure 3.10 displays the results for a search of any websites from 2013 with the phrase Michael Bazzell and the word OSINT while excluding the word Mike.

Search Engine Colossus (searchenginecolossus.com)

This website is an index of practically every search engine in every country. The main page offers a list of countries alphabetically. Each of these links connects to a list of active search engines in that country. I stay away from this service when searching American-based subjects. However, if

my target has strong ties to a specific country, I always research the engines that are used in that area through this website.

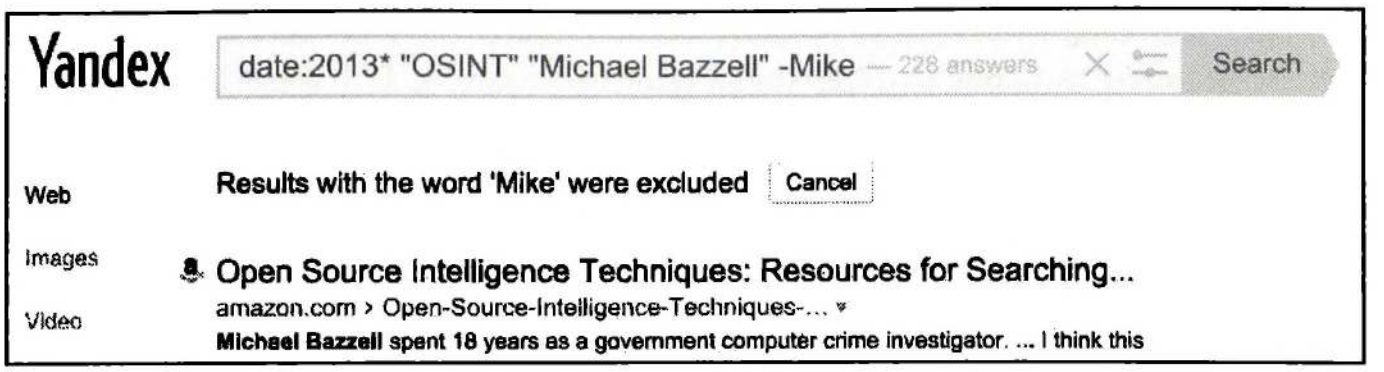


Figure 3.10: A custom Yandex search.

Duck Duck Go (duckduckgo.com)

This search engine with a clean interface offers two unique services. It has gained a lot of popularity because it does not track anything from users. Engines, such as Google, record and maintain all of your search history and sites visited. This can be a concern to privacy advocates and those with sensitive investigations. Additionally, it uses information from crowd sourced websites such as Wikipedia and Wolfram Alpha to augment traditional results and improve relevance. You will receive fewer results here than at more popular search engines, but the accuracy of the results will improve.

Start Page (startpage.com)

Similar to Duck Duck Go, Start Page is a privacy focused search engine that does not reveal your connection information to traditional search engines. The difference here is that Start Page only includes Google results versus Duck Duck Go's collaboration of multiple sources. The benefit to this is the ability to use Google's advanced search options while still protecting your identity. This includes filtering by date, images, and videos. Another benefit is the ability to open any result through a "proxy" link. This option, labeled "Proxy" next to each search result, opens the linked page through Start Page's servers and displays the content within their site. This protects your IP address from anyone monitoring connections at the target website. While this technique is not foolproof, it provides a valid layer of protection. My search strategy involves Start Page whenever I have a sensitive search that I do not want to associate with my computer or internet connection. This might include investigations that involve highly sensitive topics such as tech savvy stalker suspects.

FTP Search

I believe that the searching of File Transfer Protocol (FTP) servers is one of the biggest areas of the internet that is missed by most online researchers. FTP servers are computers with a public IP address used to store files. While these can be secured with mandated access credentials, this

is rarely the case. Most are public and can be accessed from within a web browser. The Overall use of FTP to transfer files is minimal compared to a decade ago, but the servers still exist in abundance. I prefer the manual method of searching Google for FTP information. As mentioned earlier, Google and Bing index most publicly available data on FTP servers. A custom search string will be required in order to filter unwanted information. If I were looking for any files including the term "confidential" in the title, I would conduct the following search on Google and Bing.

inurl:ftp -inurl:(http|https) "confidential"

The result will include only files from ftp servers (inurl:ftp); will exclude any web pages (-inurl: (http|https); and mandate that the term "confidential" is present (""). I have located many sensitive documents from target companies with this query. The above search yielded 107,000 FTP results. However, these specific hits are not the only valuable data to pursue. Consider the following example. I want to locate PDF documents stored on FTP servers that contain "cisco" within the title or content, and I conduct the following search on Google.

inurl:ftp -inurl:(http|https) "cisco" filetype:pdf

This results in 20,000 options within multiple FTP servers hosted on numerous domains. The first result is hosted on the Southwest Cyberport FTP server and connects to a PDF document at the following address. It appears to be a chapter of a textbook.

ftp://ftp.swcp.com/pub/cisco/03chap01.pdf

Manually changing the last "01" to "02" loads the second chapter of the book. However, it is easier to eliminate the document name altogether and browse the directory titled "cisco". The first of the following addresses displays the contents of that folder, while the second displays the content of the "pub" folder. Copy these directly into a web browser to see the results.

ftp://ftp.swcp.com/pub/cisco/
ftp://ftp.swcp.com/pub/

This type of manual navigation will often reveal numerous publicly available documents that traditional searches withhold. I have located extremely sensitive files hosted by companies, government agencies, and the military. Most File Transfer Protocol (FTP) servers have been indexed by Google, but there are other third party options that are worth exploring. At the end of each description, I identify the number of results included for the search "Cisco" "PDF".

Global File Search (globalfilesearch.com)

Global File Search provides one of the few web-based engines for searching files on these public servers. At the time of this writing, the site claims to have indexed 243 terabytes of files in public

FTP servers. Anyone searching for intelligence on any business should take a look at this site. The results are usually minimal, but very reliable.

"Cisco" "PDF": 121

File Mare (filemare.com)

Although it has indexed a smaller amount of data than the others, File Mare should be visited to locate possible FTP content. Much of the data on this website is no longer available. However, knowing the IP address of a previous FTP server with desired content could be valuable to your investigation. Navigating to that public address may reveal new documents that have yet to be indexed.

"Cisco" "PDF": 5,739

Napalm FTP (searchftps.org)

This FTP search engine often provides content that is very recent. After each result, it displays the date that the data was last confirmed at the disclosed location. This can help locate relevant information that is still present on a server. While it generated the most results of all four services, many of them were no longer available on the target FTP servers. Some could be reconstructed with cached copies, but not all.

"Cisco" "PDF": 7,947

Mamont (mmnt.ru)

This Russian FTP server allows you to isolate search results by the country that is hosting the content. This is likely determined by IP address. While most of the filtered results will be accurate, I recommend searching through the global results before dismissing any foreign options. My favorite feature of this engine is the "Search within results" option. After conducting my search, I checked this option and my search field was cleared. I entered "router" and clicked search again. I was prompted with the 436 results within my original hits that also included the word router. While this could have been replicated manually, I appreciate the option.

"Cisco" "PDF": 4,673

For comparison, Google found 23,600 results for inurl:ftp -inurl:(http|https) "Cisco" "PDF".

Nerdy Data (nerdydata.com/search)

Google, Bing, and other search engines search the content of websites. They focus on the data that is visually present within a web page. Nerdy Data searches the programming code of a

website. This code is often not visible to the end user and exists within the HTML code, JavaScript, and CSS files with which most users are not familiar. This code can be extremely valuable to research in some scenarios. Viewing the source code of a website can be done by right-clicking on the background of a page and selecting "View Source". The following two examples should explain a small portion of the possibilities with this service.

In Chapter Sixteen, you will learn about free services that try to identify additional websites that may be associated with your target website. The backbone of these services relies on the indexing of programming data of websites. Nerdy Data may be the purest way of searching for this data. If you were to look at the source code of one of my previous websites (no longer online), you would have seen at the bottom that I used a service called Google Analytics. This service identifies the number of visitors to a website and the general area where they are located. The following is the actual code that was present.

```
<script type="text/javascript">
try {var pageTracker = _gat._getTracker("UA-8231004-3");
pageTracker._trackPageview();
} catch(err) {}</script>
```

The important data here is the "UA-8231004-3". That was my unique number for Google Analytics. Any website that I used the service with would have needed to have that number within the source code of the page. If you searched that number on Nerdy Data, you will get interesting results. Nerdy Data previously identified three websites that were using that number, including computercrimeinfo.com and two additional sites that I maintain for a law firm. You can often find valuable information within the source code of your target's website.

Many web designers and programmers steal code from other websites. In the past, this would be very difficult to identify without already knowing the suspect website. With Nerdy Data, you can perform a search of the code of concern and identify websites that possess the data within their own source code. In 2013, I located a custom search website at the YGN Ethical Hacker Group that inspired me to create my own similar search service. I was curious if there were any other search websites that possessed this basic code that might give me more ideas. I looked at the source code of the website and located a small piece of code that appeared fairly unique to that service. I conducted a search on Nerdy Data for the following code.

```
<li>http://yehg.net/q?[keyword]&c=[category]  (q?yehg.net&c=Recon)</li>
```

This code was within the JavaScript programming of the search website. The search results identified 13 websites that also possessed the same code. Two of these results were hosted on the creator's website, and offered no additional information. Three of the results linked to pages that were no longer available. Three of the results linked to pages that were only discussing the code within the target website and how to improve the functionality. However, four of the results identified similar search services that were also using the programming code searched. This

revealed new search services that were related to the website in which I was interested. This same technique could be used to identify websites that are stealing proprietary code, locate pages that were created to attempt to fool a victim into using a cloned site, or validate the popularity of a specific programming function being used on hacking websites globally.

Qwant (qwant.com)

Qwant attempts to combine the results of several types of search engines into one page. It was launched in 2013 after two years of research. It has an easily digestible interface that displays results in columns titled Web, News, Knowledge, Social, and Shopping. There is a Google "feel" to it and the layout can be changed to your own preferences. A default search of my own name provided the expected results similar to Google and Bing. Clicking on the "People" tab at the top introduced new results not found on the other engines. The results included recent posts from Twitter, Facebook, LinkedIn, and MySpace from and about people with my name.

IntelTechniques Search Tool (inteltechniques.com/OSINT/user.html)

At this point, you may be overwhelmed with the abundance of search options. I can relate to that, and I do not take advantage of every option during every investigation. During my initial search of a target, I like to rely on the basics. I first search Google, Bing, Yandex, and the smaller Google search engines. In order to assist with this initial search, I created a custom tool that will allow you to quickly get to the basics. Figure 3.11 displays the current state of this online website. The main set of search options will allow you to individually search directly through Google, Bing, Yahoo, Searx, Yandex, Baidu, Exalead, DuckDuckGo, Start Page, Google Newsgroups, Google Blogs, FTP Servers, data folders, Google Scholar, Google Patents, Google News, Google Newspapers, The Wayback Machine, and Qwant. After that, the Tor options discussed previously are available. The last section presents the Newspaper Archive search options that were explained earlier. Across all options, each search that you conduct will open within a new tab within your browser. This tool is based on simple JavaScript and your queries are never seen nor stored on my server. The search all takes place on your computer within your browser, directly to the sources.

The "Submit All" option will allow you to provide any search term that will be searched across all of the services listed. Each service will populate the results within a new tab in your internet browser. Regardless of the browser that you use, you must allow pop-ups for my domain in order for the tool to work. You can also use any of the search operators discussed previously within this tool.

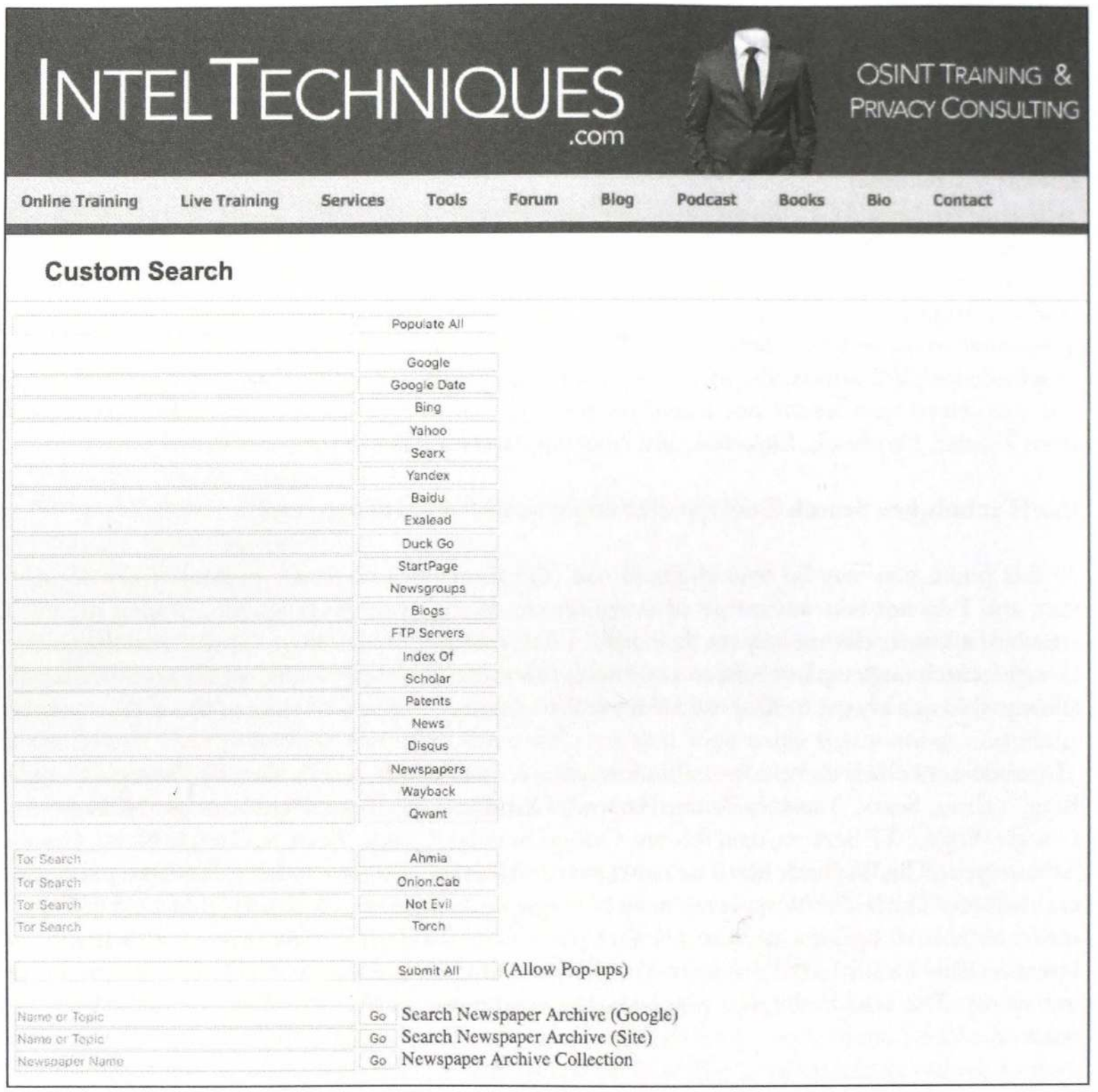


Figure 3.11: The IntelTechniques Custom Search Engines Tool.

CHAPTER FOUR
SOCIAL NETWORKS: FACEBOOK

There are hundreds of social networks that act as storage for details of a person's life. Information that was once held privately within a small group of friends or family is now broadcasted to the world via public websites. Searching these websites has always been a high priority for intelligence gathering. This chapter should identify some new techniques that can be applied on any target. Before proceeding with any of the methods here, it is important to discuss covert accounts.

Some networks' search options are severely limited without being logged into an account. In fact, I do not recommend trying any searches on Facebook without having a clean account in place. Covert accounts on all of the social networks mentioned here are free and can be completed using fictitious information. However, some networks will make this task more difficult than others. While services such as Google will give anyone multiple Gmail and Google+ accounts with little verification, Facebook, Twitter, Instagram, and Yahoo are known to make you jump through hoops before you are granted access. We begin this chapter discussing ways around these roadblocks.

Email: It is vital that you possess a "clean" email address for your covert accounts. Every social network requires an email address as part of account registration, and you should never use an already established address. Chapter Eight explains methods for researching the owners behind email addresses, and those techniques can be applied to you and your own accounts. Therefore, consider starting fresh with a brand-new email account dedicated toward use for covert profiles.

The choice of email provider is key here. I do not recommend Gmail, Yahoo, MSN, or any other extremely popular providers. These are heavily used by spammers and scammers, and are therefore more scrutinized than smaller providers. My preference is to create a free email account at GMX (gmx.com). This established mail provider is unique in two ways. First, they are one of the only remaining providers that do not require an established email address in order to obtain a new address. This means that there will be no connection from your new covert account to any personal accounts. Second, they are fairly "off-radar" from big services such as Facebook, and are not scrutinized for malicious activity. GMX will provide anyone unlimited free accounts. I suggest choosing an email address that ends in gmx.us instead of gmx.com, as that domain is less used than their official address. This is a choice during account creation. Once you have your new email address activated, you are ready to create covert profiles.

Facebook: This is by far the most difficult in terms of new account creation. For most new users, Facebook will require you to provide a cellular telephone number where a verification text can be sent and confirmed. Providing VOIP numbers such as a Google Voice account will not work anymore. I have found only one solution. Turn off any VPN, Tor Browser, or other IP address masking service and connect from a residential or business internet connection. Make sure that

you have run CCleaner to clear out all of your internet cache and log you out of any accounts. Instead of creating a new account on facebook.com, navigate directly to m.facebook.com. This is the mobile version of their site which is more forgiving on new accounts. During account creation, provide the GMX email address that you created previously. In most situations, you should bypass the requirement to provide a cellular number. If this method failed, there is something about your computer or connection that is making Facebook unhappy. Persistence will always equal success eventually.

Twitter: Many of the Twitter techniques presented here will not require an account. However, the third-party solutions will mandate that you be logged into Twitter when using them. I highly recommend possessing a covert account before proceeding. As long as you provide a legitimate GMX email address from a residential or business internet connection, you should have no issues. You may get away with using a VPN to create an account, but not always.

Instagram: Instagram is similar to Twitter. Unless you are creating multiple accounts per day from the same internet connection, you should receive no resistance in creating an anonymous account.

Yahoo: A Yahoo email account will be required later for an optional Facebook search technique. Unfortunately, Yahoo is very protective of new accounts. Most new users are mandated to provide a current cellular telephone number that can be verified. Occasionally, providing a lesser known email address (GMX) during registration from an IP address that has not created new accounts in the past 30 days will get you past this hurdle. The success of this has been sporadic. If you meet resistance from Yahoo while creating a new account, my advice is to skip them. It will not be vital to your research.

Google/Gmail/Voice: While Google has become more aggressive at refusing suspicious account registrations, they are still very achievable. As with the previous methods, Google will likely block any new accounts that are created over Tor or a VPN. Providing your GMX address as an alternative form of contact during the account creation process usually satisfies their need to validate your request. I have also found that they seem more accommodating during account creation if you are connected through a standard Chrome browser versus a privacy-customized Firefox browser. While that seems a bit shady, it makes sense (Google owns Chrome).

Some readers may assume that they can simply use their personal and accurate social network account to search for information. While this is indeed possible, it is risky. Some services, such as Instagram may never indicate to the target that your specific profile was used for searching. Others, such as Facebook, will indeed eventually notify the target that you have an interest in him or her. This is usually in the form of friend recommendations. On any service, you are always one accidental click away from sending a friend request from your real account to the suspect. For these reasons, I never use a real social network profile during any investigation. I like to maintain multiple accounts at all times in case one is suspended or deleted by the social network.

The topic of undercover operations quickly exceeds the scope of this book about search techniques. Volumes could be written about proper photo use and the psychology of posts in order to create an assumption that the person is real. For our purposes, we only need a legitimate account. We will not add any personal details or photos. We will not post any messages privately or publicly. We simply need to be logged into real accounts in order to pacify the social networks. I will assume that you now have social network accounts created that are not in your real name and possess no personal identification about you. It is time to dig into social network profiles and extract data.

Facebook users tend to keep their information a little more secure than users of other social networking sites. By default, a new Facebook user must specify the privacy settings to their account during the creation of their profile. This is mostly thanks to privacy advocates that continuously protest Facebook's privacy policies. Many of these user settings simply do not promote privacy and leave the user's information exposed for anyone to see. This section will explain numerous ways to obtain user information that is not visible on the public profile.

Once logged in, a simple search field will be present at the top of any Facebook page. Typing in the target's real name should lead to some results. Unlike Twitter, Facebook users usually use their real name when creating a profile. This profile is also usually linked to an employer, graduating high school class, or college alumni. Once a user's profile is located, the default view is the "timeline" tab. This will include basic information such as gender, location, family members, friends, relationship status, interests, education, and work background. This page will also commonly have a photo of the user and any recent posts on their page. With billions of active users, it will be likely that you will locate several user profiles under the same name as your target. There are a few things that you can do to find the right person.

If your target's name is Tom Johnson, you have your work cut out for you. This does not mean that you will never find his Facebook page, but you will need to take additional steps to get to your target. When searching the name, several possibilities will appear in a drop-down menu. This is obviously not the complete list of Tom Johnsons that are present on Facebook. At the bottom of this list is an option to see all of the profiles with your target name. After scrolling down through this list, you can select "See more results" to continue loading profiles with your target's name. You can look through these and hope to identify your target based on the photo, location, or additional information displayed in this view.

In 2013, Facebook introduced the Graph search, which eliminated some of the search filters visible on the people search pages. These filters were still available throughout 2014, but you needed to enter specific words into the search form to use them. This was often achieved by common phrases such as "people named Tom Johnson who live in Chicago, Illinois". If your wording was perfect, you would receive accurate results. In February of 2015, Facebook removed the ability to conduct detailed searches for user profiles within the search field. Prior to this time, you could enter the following queries directly into a Facebook page and receive matching results.

People named "John Smith"
People who work at Microsoft
People who work in Chicago
People who live in Chicago
People who like OSINT
People who attended Arlington High School
People born in 1980
People who visited Peru
People who speak Russian

Many of these searches still function, but you will not receive all of the results you should. You could also combine searches via keyword and obtain filtered content. Prior to 2015, a search of people who work at Microsoft and live in Seattle and like Starbucks would return the appropriate results. Today, that search produces an error. However, we can recreate these types of searches and receive more detailed information than ever before. Facebook now requires a specific address (URL) in order to search this type of information. It is not as user friendly as their previous option, but we can recreate each specific query. The following section will identify the basic types of information you can search and the necessary detailed structure of each address. After, I will explain a more sophisticated process that will help hone in on specific targets. At the end, I will present my custom online Facebook search tool that will simplify the entire process. If any of this seems overwhelming, know that my online tool will automate all of the searches.

Name

Facebook still offers native name searching within any page. However, I have received better results by using a custom address. The following URL would identify any profiles with my name. Note the "%20" in the URL represents a space. Most browsers will automatically convert a space to this format. It can be replicated by typing "People named Michael Bazzell" in the search field.

https://www.facebook.com/search/str/Michael%20Bazzell/users-named

Employment

Facebook can help you find current and past employees of a business. This can be beneficial if you are investigating a shady company and would like to contact employees that can assist with your investigation. I have also used this to gather intelligence from employees that were committing illegal acts while at their workplace. Knowing a suspect's friends with whom they work can be valuable during an interview. Knowing the names of people that work with your suspect, but are not necessarily friends with the suspect, can be more valuable. I might focus on these people as potential witnesses since they probably do not have as much loyalty to the suspect as one of the target's friends. If I wanted to locate all Facebook profiles of Microsoft employees, the following address would provide the results. As of this writing, searching "People who work at Microsoft" was also successful.

https://www.facebook.com/search/str/Microsoft/pages-named/employees/present

Changing "present" to "past" within this address will identify profiles of people that are likely no longer employed by the target company. You can also combine search options into one URL. You need to place each search structure into a single address and add "intersect" at the end. If you wanted to search for all previous Microsoft employees named "Mike Smith", the following address would produce results.

https://www.facebook.com/search/str/Mike%20Smith/users-named/str/Microsoft/pages-named/employees/past/intersect

Location

Searching profiles by location can be extremely beneficial when you do not know the name of your target. If you know the city where your target lives, you can use this as a filter. The address of a search to identify every user that currently lives in Denver is as follows.

https://www.facebook.com/search/str/denver/pages-named/residents/present

This will reveal an enormous amount of profiles. You can filter them by adding more search criteria. The following URL would display only users that live in Denver and are chefs.

https://www.facebook.com/search/str/Denver/pages-named/residents/present/str/chef/pages-named/employees/present/intersect

If you know that your target's name is Ellen, and she lives in Denver, and she is a chef, the following URL will identify the one person on Facebook that matches this search. Again, notice the ability to change present to past for previous results.

https://www.facebook.com/search/str/Denver/pages-named/residents/present/str/chef/pages-named/employees/present/str/Ellen/users-named/intersect

Likes

Facebook will allow you to filter results by a topic of interest. When people click the "like" button on a Facebook page, their profiles can be searched by this data. The URL to view every Facebook user that likes Budweiser would be as follows.

https://www.facebook.com/search/str/budweiser/pages-named/likers

I was once asked to look into an incident at a community college where vulgar graffiti was found next to a spray-painted logo of a specific local musical group. I searched for people that liked this band on Facebook and attended the target high school. This produced a list of four subjects that

attended the school and were fans of the band. All four subjects were questioned and one of them confessed to being present at the scene and identified the culprit. If I were searching for Harvard students that like the band Thrice, the following address would produce the results.

https://www.facebook.com/search/str/Harvard/pages-named/students/str/Thrice/pages-named/likers/intersect

Education

Facebook will also allow you to search for students of a specific school without knowing names. This can help create a list of potential witnesses to an event and will allow you to analyze friendships. The following search would identify current and past students of Harvard University:

https://www.facebook.com/search/str/harvard/pages-named/students

Age

I have found age range searching beneficial when my target is using a false name on Facebook or an alias that I do not know. This has also been used to identify sex offenders that provide real information, but change their name on Facebook to avoid detection. Choosing the age range and city of residence of your target may provide any Facebook pages that fit the criteria. You may receive several results that do not apply to your subject, but finding your target within these pages can be very gratifying. The following URL would identify people named Brad O'Neal between the ages of 37 and 40.

https://www.facebook.com/search/str/Brad%20O'Neal/users-named/str/37/40/users-age-2/intersect

If you know the exact year of birth of your target, you can also search by this data. If you know that your target's name is Tim Smith, and that he was born in 1980, the following address would display the appropriate results. This filtered numerous people with a common name to only seven profiles.

https://www.facebook.com/search/str/Tim%20Smith/users-named/str/1980/date/users-born/intersect

Beginning in 2016, I have found this technique to only work sporadically. While we should understand the operation and attempt when it may be useful, do not rely on this for complete results. As with any methods on Facebook, never assume that a lack of results equals a lack of profile for your target. It just means that you need to keep tweaking the options until your subject appears.

Visited

Facebook allows users to "check in" to places that they visit. Searching this data can help identify people that were present at a specific location or event. The address to search every user that visited Wrigley Field is the following. You could easily add additional filters to identify a target.

https://www.facebook.com/search/str/wrigley%20field/pages-named/visitors

Language

Many people identify the languages that they speak within their Facebook profile. You can search this content with a query similar to the previous. The URL to identify every user that speaks Japanese and lives in Denver would be the following.

https://www.facebook.com/search/str/Japanese/pages-named/speakers/str/Denver/pages-named/residents/present/intersect

Gender

When you do not know your target's name you will likely use the previous techniques to filter the profiles. Filtering further by gender can often remove half of the possibilities. In the past, you could type something similar to "women that work at Microsoft" directly into the search field. Today, that no longer works. However, you can add "/males" or "/females" near the end of every previous address to filter by gender. The following address would display every Facebook profile of female Microsoft employees that live in Seattle. The final section of my custom Facebook search tool, which is explained later, allows you to easily specify the gender of users during your searches.

https://www.facebook.com/search/str/seattle/pages-named/residents/present/str/microsoft/pages-named/employees/present/females/intersect

Dating Interests

Many Facebook users indicate whether they are interested in meeting men or women. This can commonly be seen on their profile, and we can also search by this indicator. The following URLs would identify women in interested in men (first), men interested in women (second), men interested in men (third), and women interested in women (fourth).

https://www.facebook.com/search/females/males/users-interested/intersect/
https://www.facebook.com/search/males/females/users-interested/intersect/
https://www.facebook.com/search/males/males/users-interested/intersect/
https://www.facebook.com/search/females/females/users-interested/intersect/

This search alone would be very unhelpful. However, we can combine this option with one of our previous searches to focus on this element. The following URL would identify women who work at Microsoft that are interested in meeting men.

https://www.facebook.com/search/str/Microsoft/pages-named/employees/present/females/males/users-interested/intersect/

Posts

In the first edition, I recommended third party websites to search Facebook post content. Some of these sites still exist, but most have either shut down or stopped functioning. In order to present a more permanent solution, I have identified a reliable way to do this through Facebook. While logged into a Facebook account, one should be able to natively search within the content of public posts. However, at the time of this writing, the basic search option was not functioning as reliably as it should. A search of the term "OSINT" in the Facebook search field identified some Facebook posts that contain the term within the content. A search of Posts about "OSINT" also produces results, but neither method identified all possible public posts. The reasons that this search does not completely work are unknown. However, we do know that Facebook does support this type of search.

Instead of searching from the search field, you should conduct your query through the URL. You can successfully perform the previously failed search by entering the following address directly into your browser. Notice that the search field is identical in each example, but only the direct URL retrieved meaningful results.

https://www.facebook.com/search/str/OSINT/stories-keyword

Real World Application: A large company with thousands of employees wanted to see if any employees were online discussing sensitive information about a stealth product. Several searches on specific employee profiles returned no results. A search on Facebook of the name of the secret project identified two employees leaking inappropriate information.

Profile Information

At this point, you should be able to locate a target's profile, analyze the publicly available content, and search by topic. That is just the tip of the iceberg. Facebook collects a lot of additional information from everyone's activity on the social network. Every time someone "Likes" something or is tagged in a photo, Facebook stores that information. Until recently, this was sometimes difficult to locate, if not impossible. You will not always find it on the target's profile page, but the new Facebook Graph search allows us to dig into this information.

The official way to search this data on Facebook has flaws. You can type in what you are looking for and Facebook will give you results based on your friends and people close to your network.

As an example, when I conduct a search for photos liked by Tom Merritt, I am given no results. If I had typed "Photos liked by" and then the name of one of my friends, it would have worked fine. Since our target is not likely to be on our friends list, we must get creative in order to obtain this information.

In order to conduct the following detailed searches, you must know the user number of your target. This number is a unique identifier that will allow us to search otherwise hidden information from Facebook. Prior to mid-2015, the easiest way to identify the user number of any Facebook user was through the Graph API. While you were on a user's main profile, you could replace "www" in the address with "graph" and receive the profile ID number of that user. This no longer works because Facebook removed the ability to search their graph API by user name. However, we can still obtain this powerful number through two search options.

The first option involves viewing the source code of any user's Facebook profile. The process for this will vary by browser. In Firefox and Chrome, simply right-click on a Facebook profile page and select View Page Source. Be sure not to hover on any hyperlinks during the right-click. A new tab should open with the text only view of the source code of that individual profile. Within the browser, conduct a search on this page for "entity_id". This will identify a portion of the code within this page that contains that specific term. As an example, the following is the source code immediately before and after the search result.
?entity_id=651620441&offset=3"

In this example, the user ID of this profile is 651620441. We will use this number for numerous searches within the next instruction. Alternatively, you can use my custom Facebook search tool to immediately obtain the user ID number of any profile. This tool will be explained later. Some users prefer to look at the URLs of a target's photos in order to identify the user ID, but I believe this is bad practice. If a user has no photos, this will not work. Also, Facebook's photo displays often hide this information from plain site. I prefer to rely on the source code view or my Facebook tools for this identification. This number will allow us to obtain many more details about the account. If we want to see any photos on Facebook that this subject has "liked", we can type the following address into a web browser.

https://www.facebook.com/search/651620441/photos-liked

This basic structure contains the website (facebook.com), the action (search), the user number (651620441), and the requested information (photos-liked). Since these are photos that were "liked" by the target, the results will include photos on other people's pages that would have been difficult to locate otherwise. If we had asked Facebook for this information with only the name of the target, we would have been denied. If your target has a common name, this would not work. The method described here works because we know the target's user number. There are many other options with this search. We can navigate to the following addresses to see more information about our target (user number 651620441). Explanations of each address will be explained after the list. You should replace 651620441 with the user ID of your target.

Facebook User ID Structure

Please note that the following Facebook techniques will only function if your Facebook profile language settings are set to English (US). Any other languages will produce an error. The following pages present a detailed explanation of each URL.

https://www.facebook.com/search/651620441/places-visited
https://www.facebook.com/search/651620441/recent-places-visited
https://www.facebook.com/search/651620441/places-checked-in
https://www.facebook.com/search/651620441/places-liked
https://www.facebook.com/search/651620441/pages-liked
https://www.facebook.com/search/651620441/photos-by
https://www.facebook.com/search/651620441/photos-liked
https://www.facebook.com/search/651620441/photos-of
https://www.facebook.com/search/651620441/photos-commented
https://www.facebook.com/search/651620441/photos-interacted
https://www.facebook.com/search/651620441/photos-interested
https://www.facebook.com/search/651620441/photos-recommended-for
https://www.facebook.com/search/651620441/videos
https://www.facebook.com/search/651620441/videos-by
https://www.facebook.com/search/651620441/videos-of
https://www.facebook.com/search/651620441/videos-tagged
https://www.facebook.com/search/651620441/videos-liked
https://www.facebook.com/search/651620441/videos-commented
https://www.facebook.com/search/651620441/apps-used
https://www.facebook.com/search/651620441/friends
https://www.facebook.com/search/651620441/stories-by
https://www.facebook.com/search/651620441/stories-commented
https://www.facebook.com/search/651620441/stories-tagged
https://www.facebook.com/search/651620441/stories-liked
https://www.facebook.com/search/651620441/groups
https://www.facebook.com/search/651620441/employers
https://www.facebook.com/search/651620441/employees
https://www.facebook.com/search/651620441/relatives
https://www.facebook.com/search/651620441/followers
https://www.facebook.com/search/str/651620441/events/#/date/events/intersect*
https://www.facebook.com/search/str/651620441/events-created/#/date/events/intersect*
https://www.facebook.com/search/str/651620441/events-invited/#/date/events/intersect*
https://www.facebook.com/search/str/651620441/events-joined/#/date/events/intersect*

* - You must replace the "#" in these options with a year, such as 2016, for them to function.

The "places-visited" option will display locations that the target has physically visited and allowed Facebook to collect the location information. This is often completed through a smartphone, sometimes unintentionally. This can be used to disprove alibis or verify trips.

The "recent-places-visited" option will display locations that the target has recently physically visited and allowed Facebook to collect the location information. This feature is not always reliable, and a definitive time frame of "recent" has not been established.

The "places-checked-in" option will display locations where the target has used the Facebook app to "check in". While this can be falsified, these results are usually more accurate and believable than "places-visited".

The "places-liked" option will display any physical locations that the target has clicked "like". This will often identify vacation spots, favorite bars, and special restaurants. This can be priceless information for an investigator or skip-tracer.

The "pages-liked" option will display any Facebook pages that the target clicked "like". This will often display interests of the target such as a favorite sports team, musical group, or television show. These results will include a button labeled "liked by". Clicking this will identify everyone on Facebook that liked that item. This can quickly identify the people that visit a hole-in-the-wall bar that you are investigating.

The "photos-by" option will display Facebook photos that were uploaded by the user. These will likely already be visible on the target's photos page. However, this search could potentially reveal additional images.

The "photos-liked" option was explained on the previous page. This can be beneficial if your target has a private profile. If the photos of interest are on someone else's profile that is not private, you will be able to see all of them.

The "photos-of" option will display any photos in which the target has been tagged. This search has already been proven very effective in many investigations. This will immediately locate additional photos of your target that may not be visible on the target's profile. This is helpful when the photos are private on one person's page, but not others.

The "photos-commented" option will display any photos on profiles where the target left a comment on the photo. This can be important because the target may not have "liked" the photo nor been tagged in it. The option may produce redundant results, but it should always be checked.

The "photos-interacted" option will display any photos on profiles where the target either liked or commented on the image. This may be important in the future if Facebook eliminates the other search options.

The "photos-interested" option will display any photos on profiles when Facebook believes the target has an interest. There is not much known at this point about how this is determined, but the results appear unique from any other option.

The "photos-recommended-for" option will display any photos that would likely appear on the target's home Facebook feed. This can give a glimpse into the interests of the target according to Facebook.

The "videos" option will display videos visible on the target's profile. These may or may not be directly connected to the target. They could also be videos linked to the original source with no personal ties to the subject.

The "videos-by" option will display videos that were actually uploaded by the target. These will be much more personal to the subject and will usually include more relevant content.

The "videos-of" option is similar to the "photos-of" filter. This will display videos that supposedly contain images of the target within the video itself. It could be compared to "tagging" someone inside a video.

The "videos-tagged" option is very similar to the "videos-of" filter. This will display videos which your target was tagged as being present within the video. On most executions, these will be the same as the previous option.

The "videos-liked" option will display any videos that the target clicked "like". This can also be used to establish personal interests of the target and are often of interest to parents.

The "videos-commented" option will display any videos on profiles where the target left a comment on the video. Again, this can be important because the target may not have "liked" the video nor been tagged in it. The option may produce redundant results, but it should always be checked.

The "apps-used" option will display the apps installed through Facebook. These are usually games that can be played with other people. Many of these specify the environment with which they work such as "IOS". This would indicate that the target is using an iPhone or iPad instead of an Android device.

The "friends" option should display a list of all of the target's friends on Facebook. This will be the same list visible on the main profile page. If you receive no results, the target likely has the friend's list set to "private".

The "stories-by" option will display any public posts by the target. This can often identify posts that are not currently visible on the target's profile.

The "stories-commented" option will display any public posts by any users if the target entered a comment. This is useful in identifying communication from a target with a private profile. The standard privacy options do not prevent a search of your comment history on public posts.

The "stories-tagged" option will display any posts in which the target was tagged. This tagging is usually performed because of an interest in the post.

The "stories-liked" option will display any posts that the target clicked "like". This can also be used to establish personal interests of the target.

The "groups" option will display any groups of which the target is a member. This is beneficial in identifying stronger interests of the target. In my experience, a target must only have faint interest to "like" something. However, the interest is usually strong if a group related to the topic is joined by the target.

The "employers" option will display a list of current and previous employers of the target, as defined by the target.

The "employees" option will display profiles of people that claim to be employed by the same employer as the target. This is a quick way to identify potential co-workers. This will require the target to include an employer or previous employer to function properly.

The "relatives" option will display a list of people that the target has identified as a relative. Often, this will display relatives even if the target has the friend's list set to "private".

The "followers" option will display a list of people that follow the target on Facebook. This is usually reserved for celebrities and those with a public life, but everyday citizens are now also starting to embrace this culture.

The "events" option will display any Facebook events associated with your target. These often include parties, company events, concerts, and other social gatherings. This will almost always display events that are not listed on the target's public profile. Note that you must include a year of interest for all event options to populate. Facebook has disabled most "catch-all" queries.

The "events-created" option will display only the Facebook events created by the target. This could identify events where your target was the host or organizer.

The "events-invited" option will display any events which your target was invited to attend by another Facebook user. An invite does not necessarily indicate attendance.

The "events-joined" option will only display the Facebook events where the target acknowledged attendance. This could be in the form of a "R.S.V.P." or confirmation by the target that he or she is currently at the event.

Friends' Information

Occasionally, you may find that your target's Facebook profile possesses minimal information. The techniques mentioned previously may not locate any valuable information if your target does not provide any content to his or her profile. Therefore, knowing the overall interests of someone's friends on Facebook may assist with your analysis of the target. The following direct addresses focus only on the friends of an individual. Explanations of each URL will follow the list. Similar to the previous example, a user number is already included in the address (651620441). You should replace this with the user number of your target.

https://www.facebook.com/search/651620441/friends/places-visited
https://www.facebook.com/search/651620441/friends/recent-places-visited
https://www.facebook.com/search/651620441/friends/places-checked-in
https://www.facebook.com/search/651620441/friends/places-liked
https://www.facebook.com/search/651620441/friends/pages-liked
https://www.facebook.com/search/651620441/friends/photos-by
https://www.facebook.com/search/651620441/friends/photos-liked
https://www.facebook.com/search/651620441/friends/photos-of
https://www.facebook.com/search/651620441/friends/photos-commented
https://www.facebook.com/search/651620441/friends/photos-interacted
https://www.facebook.com/search/651620441/friends/photos-interested
https://www.facebook.com/search/651620441/friends/photos-recommended-for
https://www.facebook.com/search/651620441/friends/videos
https://www.facebook.com/search/651620441/friends/videos-by
https://www.facebook.com/search/651620441/friends/videos-of
https://www.facebook.com/search/651620441/friends/videos-tagged
https://www.facebook.com/search/651620441/friends/videos-liked
https://www.facebook.com/search/651620441/friends/videos-commented
https://www.facebook.com/search/651620441/friends/apps-used
https://www.facebook.com/search/651620441/friends/friends
https://www.facebook.com/search/651620441/friends/stories-by
https://www.facebook.com/search/651620441/friends/stories-commented
https://www.facebook.com/search/651620441/friends/stories-tagged
https://www.facebook.com/search/651620441/friends/stories-liked
https://www.facebook.com/search/651620441/friends/groups
https://www.facebook.com/search/651620441/friends/employees
https://www.facebook.com/search/651620441/friends/employers
https://www.facebook.com/search/651620441/friends/relatives
https://www.facebook.com/search/651620441/friends/followers
https://www.facebook.com/search/651620441/friends/events
https://www.facebook.com/search/651620441/friends/events-created
https://www.facebook.com/search/651620441/friends/events-invited
https://www.facebook.com/search/651620441/friends/events-joined

The "friends/places-visited" option will display the common places that your target's friends visit. This may identify local hangouts and favorite bars.

The "friends/recent-places-visited" option will display the common places that your target's friends have recently visited. This may identify more relevant places of interest.

The "friends/places-checked-in" option will display the common places that your target's friends have checked-in through the Facebook app or website. This may identify more accurate places of interest.

The "friends/places-liked" option will display the common places that your target's friends like on Facebook. This can also identify locations of interest to the investigator.

The "friends/pages-liked" option will display the common pages that your target's friends like on Facebook. This will usually identify products and items of interest instead of places. I have used this to determine if my target's friends were mostly interested in drugs or other criminal activity.

The "friends/photos-by" option will display the photos uploaded by your target's friends. This can quickly identify interests and recent locations visited by the subjects. Occasionally, you may locate untagged photos of your target in this collection.

The "friends/photos-liked" option will display the photos "liked" by your target's friends. This can provide an indication of interests by both the target and the friends.

The "friends/photos-of" option will display the tagged photos of your target's friends. This will quickly display images which may identify recent locations visited by the friends.

The "friends/photos-commented" option will display the photos on Facebook where your target's friends placed a comment. This often includes communication otherwise difficult to find.

The "friends/photos-interacted" option will display the photos on Facebook where your target's friends liked or commented on a photo. This combines two previous efforts.

The "friends/photos-interested" option will display the photos which Facebook determined were of interest to your target's friends.

The "friends/photos-recommended-for" option will display the photos on Facebook which would likely be on the overall profile feed of your target's friends.

The "friends/videos" option will display the videos on your target's friends' pages.

The "friends/videos-by" option will display the videos uploaded by your target's friends.

The "friends/videos-of" option will display the videos containing your target's friends. This may include video taken at events where your target was present.

The "friends/videos-tagged" option will display the videos tagged for your target's friends. This will be very similar to the previous option.

The "friends/videos-liked" option will display the videos "liked" by your target's friends. This can provide an indication of interests by both the target and the friends.

The "friends/videos-commented" option will display the videos on Facebook where your target's friends placed a comment. This often includes communication otherwise difficult to find.

The "friends/apps-used" option will display the apps that your target's friends use on Facebook.

The "friends/friends" option will display the friends of your target's friends on Facebook.

The "friends/stories-by" option will display the Facebook posts that your target's friends posted.

The "friends/stories-commented" option will display the common Facebook posts that contain comments from your target's friends. This can often display valuable communication that would otherwise be very difficult to locate.

The "friends/stories-tagged" option will display the common Facebook posts that your target's friends tagged as interesting.

The "friends/stories-liked" option will display the overall Facebook posts which your target's friends clicked the "like" button while viewing.

The "friends/groups" option will display the Facebook groups that contain the most members of your target's friends. This is another valuable source for locating possible interests of your target.

The "friends/employees" option will display the profiles of users that work at the same place as your target's friends. This can lead to the discovery of cliques associated with the target but not necessarily listed within the target's profile.

The "friends/employers" option will display the companies which your Facebook target's friends are currently or were previously employed. This can lead to the discovery of common businesses associated with the target but not necessarily listed within the target's profile.

The "friends/relatives" option will display the relatives of your target's Facebook friends. I once found value in this when trying to locate people that may cooperate with an investigation. During a homicide investigation, many of the people involved refuse to talk to the police. However,

reaching out to relatives often provides more support. Additionally, repeating this search with the user number of identified relatives will often display additional unlisted relatives.

The "friends/events" option will display the common events with which your target's friends have an association. This may identify upcoming large local gatherings. I have successfully used this information to conduct surveillance for a wanted fugitive.

The "friends/events-created" option will display the common events that your target's friends created.

The "friends/events-invited" option will display the common events with which your target's friends were invited.

The "friends/events-joined" option will display the common events that your target's friends attended.

Common Results

The previous options will easily allow you to view sensitive information about an individual target. Repeating these queries on additional individuals may allow you to recognize common patterns. This may identify relationships between individuals that are not obvious through their profiles. Knowing that two people both joined a local event, liked a rare page, commented on a specific photo, or joined a small group can generate a new lead in an investigation. Much of this information would not be visible on an individual's profile. Instead of manually analyzing multiple profiles for common information, consider a specific URL address that will only display combined results.

Facebook supports this type of analysis within the standard search field. Entering "Pages liked by Mark Zuckerberg and Chris Hughes" will display numerous profiles and pages that both subjects "liked". The flaw with this type of search is that it is not unique per Facebook profile. There are dozens of profiles for people named Chris Hughes. This native search does not know which Chris Hughes you have an interest. Therefore, I never recommend this type of written query. Instead, the following URL will display only pages liked by both Mark Zuckerberg (Facebook user number 4) and Chris Hughes (Facebook user number 5).

https://www.facebook.com/search/4/pages-liked/5/pages-liked/intersect

This specific structure is important for the query to function. The URL begins with the same terms that you learned in the previous section (facebook.com/search/user number/pages-liked). It is followed by a second search for an additional user (second user number/pages-liked). The final term of "intersect" instructs Facebook to only list results that apply to both users listed. You could continue to add users to this search. The following URL will display the five Facebook pages liked by Mark Zuckerberg (Facebook user number 4), Chris Hughes (Facebook user

number 5), and Arie Hasit (Facebook user number 7).

https://www.facebook.com/search/4/pages-liked/5/pages-liked/7/pages-liked/intersect

This method could be applied to all of the techniques that were explained earlier in reference to an individual. The following URL queries would identify common elements between Mark Zuckerberg and Arie Hasit. It should be noted that I am using these examples because of the single digit user numbers. Your user number searches will likely be much longer and appear similar to "17893008278".

https://www.facebook.com/search/4/places-visited/7/places-visited/intersect
https://www.facebook.com/search/4/recent-places-visited/7/recent-places-visited/intersect
https://www.facebook.com/search/4/places-checked-in/7/places-checked-in/intersect
https://www.facebook.com/search/4/places-liked/7/places-liked/intersect
https://www.facebook.com/search/4/pages-liked/7/pages-liked/intersect
https://www.facebook.com/search/4/photos-liked/7/photos-liked/intersect
https://www.facebook.com/search/4/photos-of/7/photos-of/intersect
https://www.facebook.com/search/4/photos-commented/7/photos-commented/intersect
https://www.facebook.com/search/4/photos-interacted /7/photos- interacted /intersect
https://www.facebook.com/search/4/photos-interested /7/photos- interested /intersect
https://www.facebook.com/search/4/videos/7/videos/intersect
https://www.facebook.com/search/4/videos-of/7/videos-of/intersect
https://www.facebook.com/search/4/videos-tagged/7/videos-tagged/intersect
https://www.facebook.com/search/4/videos-liked/7/videos-liked/intersect
https://www.facebook.com/search/4/videos-commented/7/videos-commented/intersect
https://www.facebook.com/search/4/apps-used/7/apps-used/intersect
https://www.facebook.com/search/4/stories-commented/7/stories-commented/intersect
https://www.facebook.com/search/4/stories-liked/7/stories-liked/intersect
https://www.facebook.com/search/4/stories-tagged/7/stories-tagged/intersect
https://www.facebook.com/search/4/groups/7/groups/intersect
https://www.facebook.com/search/4/employees/7/employees/intersect
https://www.facebook.com/search/4/employers/7/employers/intersect
https://www.facebook.com/search/4/relatives/7/relatives/intersect
https://www.facebook.com/search/4/followers/7/followers/intersect
https://www.facebook.com/search/4/events/7/events/intersect
https://www.facebook.com/search/4/events-joined/7/events-created/intersect
https://www.facebook.com/search/4/events-joined/7/events-invited/intersect
https://www.facebook.com/search/4/events-joined/7/events-joined/intersect

Real World Application: I have found myself in interview rooms with suspects that claim to not know the co-suspects being investigated. Locating rare and specific Facebook interests or photos in common with both suspects always gave me an edge in the interview. The suspects would have a hard time explaining the coincidence. In one case, two suspects that swore they did

not know each other had both referenced going to a party together within the comment area of a long-forgotten photo on Facebook.

Common Friends

Facebook will often display all of the friends of a target. However, it does not natively display a set of friends that two users have in common. We can accomplish this with a URL trick. First, consider scenarios of when this might be useful.

A law enforcement officer that has two suspects in custody for a homicide is getting little cooperation from the duo. Identifying only the people that are friends with both of them may provide an opportunity to interview someone that has direct knowledge of the incident.

A teacher or counselor that is trying to resolve ongoing issues between two students is having a hard time getting them to talk about the situation. Casually talking to a couple of people that are friends with both of the subjects may lead to better details about the real problems.

Finally, an investigator is looking for evidence of a cheating spouse. Looking at the photos on the Facebook pages of friends of both the cheating spouse and the new boyfriend or girlfriend may provide the final piece of the investigation.

The structure of the address that will conduct this search is as follows.

https://www.facebook.com/browse/mutual_friends/?uid=USERID&node=USERID

If we wanted to display only the friends of both Arie Hasit (arie.hasit - User # 7) and Chris Hughes (ChrisHughes - User # 5), the following address would be appropriate.

https://www.facebook.com/browse/mutual_friends/?uid=7&node=5

This exact search identified 13 people of interest. More interesting is the fact that Arie Hasit's friends are private. We should not be able to see anything here, but we do. Therefore, modifying the second user name with other people that may be friends with Arie will likely display even more of these "private" friends. This link could be saved as a bookmark and you can check it often to identify anyone new that has been added to this circle of friends. Comparing a new search to a previous screen capture of this data may also identify anyone that has been removed from this group. An individual no longer accepted by his former friends may make a great candidate for an interview if this group was involved in an incident you are investigating.

You can use a similar technique to discover the length of time that two users have been friends on Facebook and immediately see commonalities of both users. The URL structure is as follows.

https://www.facebook.com/USERNAME1?and=USERNAME 2

If you wanted to see information of the two users that were mentioned earlier (Mark Zuckerberg and Chris Hughes), you would enter the following URL.

https://www.facebook.com/zuck?and=ChrisHughes

Note that these searches can use either the Facebook user names or user numbers. The example above identifies their Facebook friendship start date of November 2006, and displays that they both work at Facebook and studied at Harvard. This technique can quickly inform an investigator of the main areas in common between two Facebook users.

Facebook ID Creation Date

Digital forensics enthusiast and private investigator Josh Huff at LearnAllTheThings.net has conducted a lot of research into the assignment of user ID numbers to Facebook profiles. We know that these numbers are assigned in chronological order, but the intelligence goes much farther beyond that. His research could fill several pages, and in some situations, he can identify the month when a specific user account was created. For the sake of space and most useful details, he has provided the following to me for publication.

Facebook transitioned from 32-bit numbers, such as 554432, to 64-bit numbers that begin with 100000 between April and December of 2009. Therefore, if your target's number is less than 100000000000000, the account was likely created prior to April 2009. An account with a 15-digit ID number would have been created after December 2009. We can break down the numbers by year. The following are rough estimates that should only be used as a general guideline.

2006: Numbers less than 600400000
2007: 600400000 - 1000000000
2008: 1000000000 - 1140000000
2009: 1140000000- 100000628000000
2010: 100000629000000 - 100001610000000
2011: 100001611000000 - 100003302000000
2012: 100003303000000 - 100004977000000
2013: 100004978000000 - 100007376000000
2014: 100007377000000 - 100008760000000
2015: 100008761000000 - 100010925000000
2016: 100010926000000 - 100014946000000
2017: 100014947000000 - 100023810000000
2018: 100023811000000 -

Business Profiles

Many Facebook profiles are not attached to an individual. Instead, they are public pages associated with a business, website, or other entity. Similar to personal profiles, they are each

assigned a unique profile ID number. This number can be used for advanced searches as we did with individuals. I will start with employee searching. As explained earlier, you can simply type "people who work at Microsoft" into the search bar and obtain decent results. This works for many businesses that have a unique name, but not those that do not. As an example, consider the business Target. A search of "people who work at Target" provides many results. Within them include references to the large retail company Target, Target Photography, Target Co., Target Shooting Range, and others. A search by profile number solves this issue.

Similar to personal profiles, you can right-click on the page and select "view source code" to display the text-only code behind the scenes of a profile. A search of "entity_id" presents the user number of that page. Alternatively, you can use the custom search tool discussed in a moment for an easier solution. Entering this number within the following URL presents only the employees of the specified business. This is because we are supplying the exact user number and not a generic text search string.

https://www.facebook.com/search/str/191491890970373/employees

This same technique can be used to specify an exact user number of a specific entity's page to discover people that have other associations to a business Facebook profile. Consider the following examples using the profile ID of the Alton Police Department's user number (191491890970373). After the following URLs, I will explain the expected results.

https://www.facebook.com/search/191491890970373/visitors/intersect
https://www.facebook.com/search/191491890970373/users-checked-in/intersect
https://www.facebook.com/search/191491890970373/likers/intersect
https://www.facebook.com/search/191491890970373/photos-in/
https://www.facebook.com/search/191491890970373/videos-in/
https://www.facebook.com/search/191491890970373/stories-tagged/

The "visitors" option displays people who have claimed to visit the specified business. This is not as reliable as the next option.

The "users-checked-in" option displays users who have actually taken action to document that they were present at a location. This could include posting with geo-location features enabled.

The "likers" option lists the Facebook users that "liked" a business page. This can identify additional people of interest when you have a small organization such as a private club.

The "photos-in" option displays images that were described as being captured while at the specified business. This often uncovers user-supplied candid photos that would never be seen on the official company profile.

The "videos-in" option is similar to the previous result, but only displays videos.

The "stories-tagged" option displays Facebook posts that were tagged as being related to the target business profile. These often include complaints intended to prompt a response from the management of the organization.

Event Search

Facebook Events are a way for members to let friends know about upcoming events in their community and to organize social gatherings. Events require an event name, network, host name, event type, start time, location, and a guest list of friends invited. Events can be public or private. Private events cannot be found in searches and are by invitation only. People who have not been invited cannot view a private event's description, timeline, or photos. While Facebook currently displays an Events tab on most searches, six specific URL addresses create more precise queries that obtain better results. The following address identifies upcoming events in Alton, Illinois.

https://www.facebook.com/search/in-future/date/events/str/Alton,%20IL/pages-named/events-at/intersect/

This was based on a location keyword search of Alton, IL (%20 represents a "space"). This will capture results that have Alton and IL in the description, but may miss some that do not. If you know the exact location of an upcoming event, you can also search by the profile ID number of the associated business profile. The following address identifies all upcoming events that will be hosted at a specific bar in Alton, IL.

https://www.facebook.com/search/in-future/date/events/str/333100363450/events-at/intersect/

This can be useful to police with monitoring trouble areas, teachers with controlling school events, or parents with identifying the next big party. If neither of these display the results that you are seeking, you can attempt a generic keyword search for all upcoming events with the following web address. This results in all events that include "protest" and "police" anywhere within the description.

https://www.facebook.com/search/events/?q=protest%20police

All three of these focused on future upcoming events. You can replicate each, and switch the attention to past events that have concluded. The following addresses repeat these searches, but include only previous events by adding "in-past" to the URL.

https://www.facebook.com/search/in-past/date/events/str/Alton,%20IL/pages-named/events-at/intersect/

https://www.facebook.com/search/in-past/date/events/str/333100363450/events-at/intersect/

https://www.facebook.com/search/in-past/date/events/str/protest%20police/keywords_events/intersect/

Official Facebook Options

Every year, Facebook modifies the look and function of their pages. If the previous searches and direct URLs do not provide the exact data that you are seeking, consider using the traditional search options available on any Facebook page. Figure 4.01 displays the current filter bar at the top of every Facebook results page. This will seek Posts, Photos, Videos, Places, Groups, Events, and other options based on the provided search terms. While this is usually quite inferior to the advanced methods presented up to this point, these basics should not be excluded.

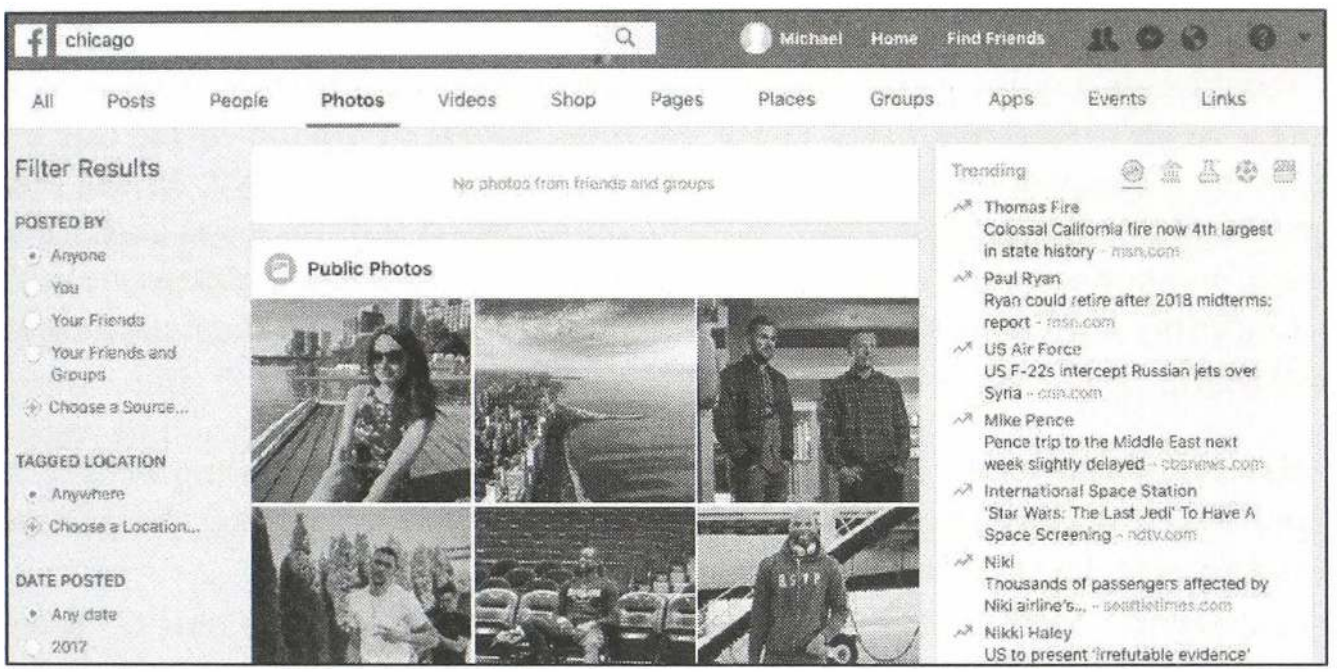


Figure 4.01: The Facebook filter options.

These filters can also be applied using direct URLs. After a search of the term "OSINT", the following addresses replicate each of the filters on a standard Facebook results page.

All: https://www.facebook.com/search/top/?q=osint
Posts: https://www.facebook.com/search/posts/?q=osint
People: https://www.facebook.com/search/people/?q=osint
Photos: https://www.facebook.com/search/photos/?q=osint
Videos: https://www.facebook.com/search/videos/?q=osint
Shop: https://www.facebook.com/search/shop/?q=osint
Pages: https://www.facebook.com/search/pages/?q=osint
Places: https://www.facebook.com/search/str/osint/keywords_places/
Groups: https://www.facebook.com/search/groups/?q=osint
Apps: https://www.facebook.com/search/apps/?q=osint
Events: https://www.facebook.com/search/events/?q=osint

Facebook Live Video Streaming

Facebook Live allows users to convert practically any smartphone into an international live video streaming device. This service has gained a lot of popularity since 2016, and there are numerous live video streams every minute. While some have thousands of viewers watching in real-time, many are never seen and lost forever. We continue to see more use of Facebook live within the execution of crimes. Two notable events were the killing of a random man in Cleveland broadcasted by the suspect and a group of teens beating a mentally challenged person in Chicago streamed by one of the offenders. I suspect that we will see a rise in this type of behavior as much of our society needs the immediate attention that this service provides. This section will focus on various research techniques that you can use during your next investigation. Note that pre-recorded Facebook videos will be discussed later in Chapter Fifteen (Videos).

First, we must locate any live video streams of interest. You may encounter a live video while researching a target's Facebook page. These will often appear at the top of the user's timeline. Clicking on the video will open the preview page, as seen in Figure 4.02. In this example, it displays the number of live viewers (680), the amount of time that the user has been broadcasting (28 minutes), and any comments from the viewers. If you have a specific target user, this is the best way to identify any live streams. If you have a target location, then the Facebook Live Map is what you need.

You can find an interactive map displaying current live streams at the following website. In Figure 4.03, the previous user's live stream is visible, and we now see that he is likely broadcasting from Georgia, but an exact location is not known yet.

https://www.facebook.com/live/discover/map



Figure 4.02: A Facebook Live Video Stream.

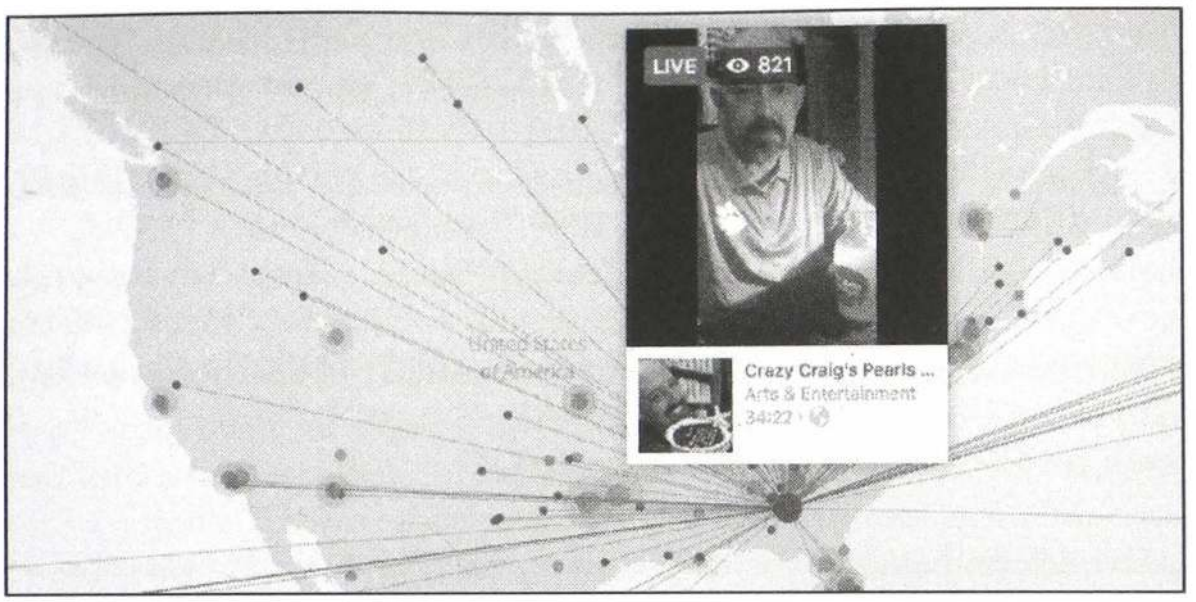


Figure 4.03: A Facebook Live Map focused on a video stream in progress.

Now that we have a target live streaming video, we will use this as our example in the following tutorials. The first piece of information that we want to uncover is a better location for the target. In Figure 4.02, we see a hyperlink under the title of his video that reads "28 minutes ago". Clicking on this link opens the target video within a standard Facebook page, instead of an embedded pop-up page, which hides the information we need. When opening that link, notice the URL of the new page. In this example, it is the following.

https://www.facebook.com/crazycraigscrafts/videos/445432539206610/

The Video ID is seen in this address, represented by "445432539206610". This video number is vital to obtain the additional information we seek. Navigate to my Facebook Search Tools as explained at the end of this chapter, and look at the lower right area. You will find a series of links similar to the following.

Video Data: ALL / Level 0 / Level 1 / Level 2

These hyperlinks each open a unique set of data from the Facebook Live servers. The links starting with "Level" connect you directly to the metadata being used to populate the Facebook Live Map that was mentioned previously. The various levels will return different details, and I provide all of them for complete coverage. The "All" option is a piece of code on my website that attempts to combine all three levels of data, remove duplicates, and return the unique results to you in a new tab. While this works most times, it may fail on occasion and I give you the "Level" alternatives if needed. For this example, I clicked the "All" option and was presented with a page full of what appears to be random text. On this page, I conducted a search within my browser by pressing command and f (Windows users would use ctrl-f) and entering the video number (445432539206610). I was immediately presented with the positive match, which appeared similar to the following.

{"videoID":"445432539206610","lat":37.487656765901,"long":-81.963124901126,"name":
"Crazy Craig's Pearls and Jewelry", "startTime":1513375622,"previewImage":
"https:\/\/scontent-lax3-1.xx.fbcdn.net\/v\/t15.0-10\/s640x640\/-24280617_-
445457842537413_3045984681699835904_n.jpg? oh=-028a7227f08-401f1c156a992-
c93dfe0d&oe-=5AD08FB7", "viewerCount" :1058,"formattedCount": "1.1K",
"publisherCategory": "Arts & Entertainment","profilePicture": "https:\/\/scontent-lax3-
1.xx.fbcdn.net\/v\/t1.0-1\/c0.0.200.200-\/p200x200\/-23795802_432652440484620_-
1993057746686495436_n.jpg?oh=bdadfed09ba6561de220b75164207b46&oe=5ACD2DA9","
width":360,"height":640,"message":"Friday", "messageRanges":[], "url":
"https:\/\/www.facebook.com\-/crazycraigscrafts\/"}

We can see in the details that this is definitely the correct video. What I am most interested in is the GPS coordinates near the beginning of the result. Note that I modified these numbers in order to better protect the privacy of the user. This GPS data represents the location of the user broadcasting the live video. Is this always accurate? No. This location can be provided due to multiple different types of data streams. In a perfect scenario, this will be a location based on the GPS identified by the mobile device being used. This happens often, but not always. If this data is not available, Facebook might use the IP address associated with the internet connection, or simply the city and state provided by the user during creation of the profile.

The accuracy of this information can vary from one extreme to another. However, this is a good lead. If that GPS location was associated with a home owned by a person with the same name, I would feel confident in the accuracy. If the location returns to the middle of the ocean, which I have seen once, then the data is obviously wrong. If your target is cautious about location sharing and uses a VPN, then these details can be quite misleading. As with everything else related to OSINT, these techniques only identify potential evidence. It is our job to validate our findings.

Next, I want to focus on the viewers of this video. Many people watch these video streams from their own mobile device, and unknowingly share their location with Facebook. We can often retrieve location information from a small percentage of the viewers, which can lead to further intelligence about the video. In our previous example, the following URL would identify location metadata about the viewers of the video stream.

https://www.facebook.com/ajax/livemap/videos/viewers/?video_id=4454325392066l0&live_viewers_count=2500&dpr=1&__user=&__a=1

The response displays location data stored within a single line of text, similar to the following.

{"latitude":51.5333,"longitude":4.46667}, {"latitude":42.9783,"longitude":-78.8},
{"latitude":36.1028,"longitude":-80.2605} {"latitude":33.7407,"longitude":-117.881},

These are the locations of some of the viewers watching the stream. As with the previous explanation about how Facebook gathers location data about its users, the same applies here.

This alone is very valuable, and we can extract these coordinates and research them on Google Maps. However, I propose a better solution.

Navigate to the Add/ Remove Line Breaks service from Text Mechanic at the following website.

http://textmechanic.com/text-tools/basic-text-tools/addremove-line-breaks/

Copy the entire data provided from the viewer location URL conducted previously. In this website, replace the text in the "Input Box" with this copied data from Facebook. In the next section is a line with a button titled "Make New Line Breaks". Select the "Before" option and place a bracket ({) in the text box following this option. This notifies the tool that we want to replace each occurrence of a "{" with a new line. Click the "Make New Line Breaks" button. This screen should now appear similar to Figure 4.04. The lower box should now contain only one location per line. If you have extra text at the beginning or end of the location data, simply delete it.

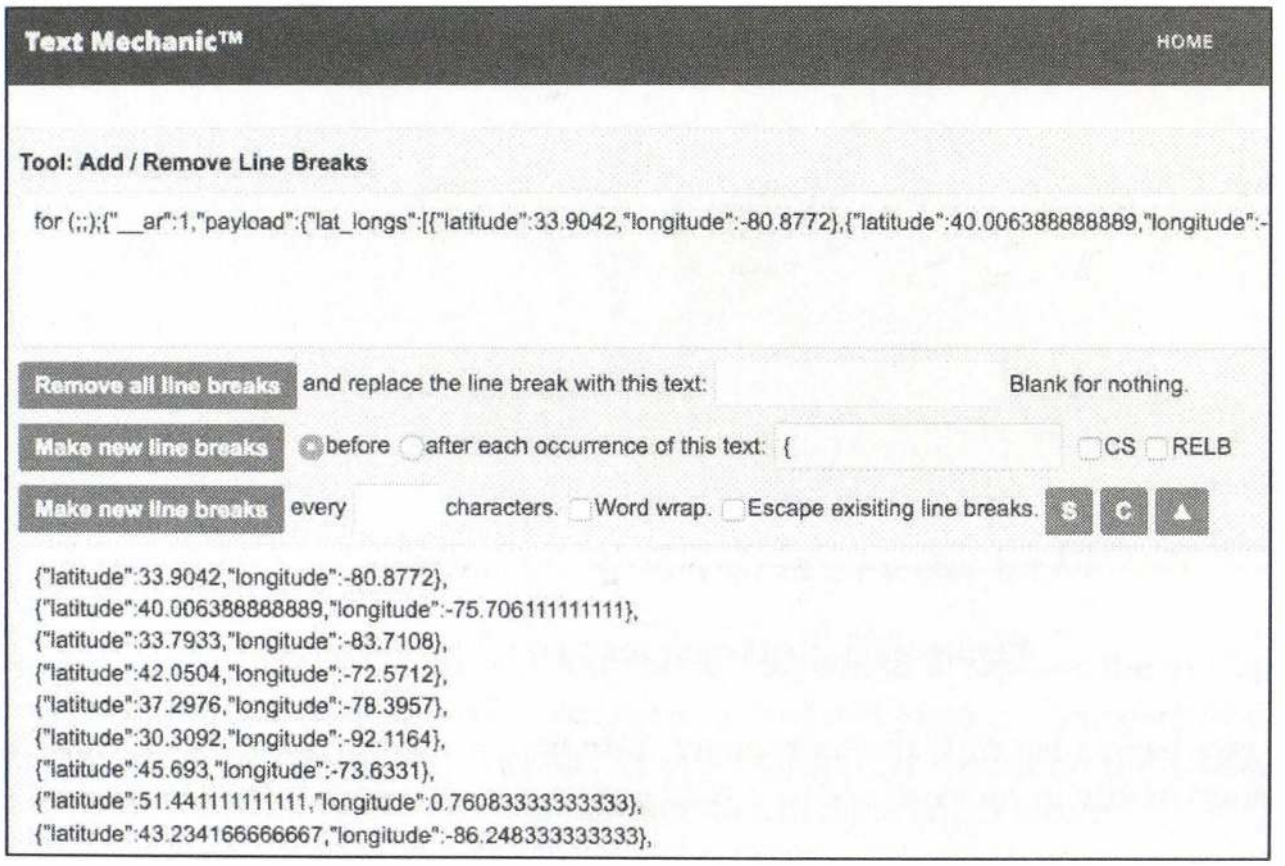


Figure 4.04: Adding line breaks to text with Text Mechanic.

Next, copy the entire text generated in the lower box of this result. Navigate to the Find and Replace Text Tool from Text Mechanic at the following link. We will use this page to clean-up our data.

http://textmechanic.com/text-tools/basic-text-tools/find-and-replace-text/

Paste the text copied previously inside the lower box that includes "Enter your text to be found and replaced here". Be sure to overwrite any text currently in that box. In the first box, titled "Find this", copy and paste the characters before the latitude of the first result from the lower box. In other words, the first box should contain {"latitude": and nothing else. Click the "Find and replace text" button and you should see that data disappear from the lower results. Repeat this by copying "longitude": from the lower box and pasting it into the first box. Click the "Find and replace text" button and you should see that data disappear from the lower results. Finally, copy and paste }, from the lower box and pasting it into the first box. Click the "Find and replace text" button and you should see that data disappear from the lower results. After this execution, your page should look similar to Figure 4.05. Notice that the lower box contains only GPS coordinates, which will be essential for the next task.

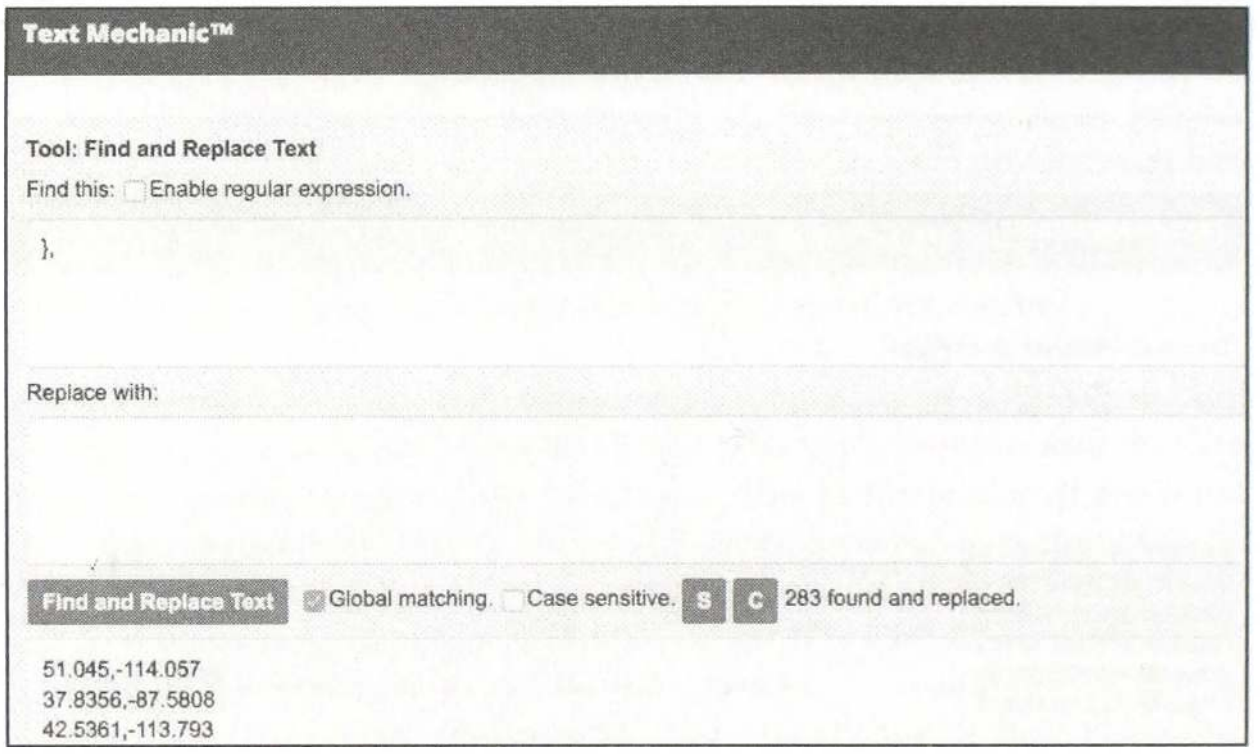


Figure 4.05: Replacing text with Text Mechanic.

Now that you have a list of GPS coordinates, it is time to identify these locations. Navigate to the following website in order to access a bulk geo-conversion tool from Doogal.

https://www.doogal.co.uk/BatchReverseGeocoding.php

Copy and paste the entire contents of the lower box from the Text Mechanic Find and Replace Text Tool to the Input box of the Doogal website. Click the "Reverse Geocode" button and witness the conversion of these coordinates into street addresses. Many of these will be the home addresses of the viewers of our target video. Below is the actual output, with modifications in order to protect the privacy of the viewers.

Latitude,Longitude,Address

33.9042,-80.8772,"58411 Lower Richland Blvd, Hopkins, SC 29061, USA"
40.006389,-75.706111,"246-38 E Pennsylvania Ave, Downingtown, PA 19335, USA"
33.7933,-83.7108,"2256 S Madison Ave, Monroe, GA 30655, USA"
42.0504,-72.5712,"969 Elizabeth Cir, Longmeadow, MA 01106, USA"
37.2976,-78.3957,"1070-1498 Pine St, Farmville, VA 23901, USA"
30.3092,-92.1164,"21739-21799 LA-93, Carencro, LA 70520, USA"

Below these results is an interactive world map that has marked each location as seen in Figure 4.06. This global view may provide a level of intelligence that was not possible from text alone. When I am researching extremist and violence-related live streams, I want to know the general locations of the viewers. Are they all local and ready to burn down the city? That presents a very specific threat. Are they all within a country ŏn our radar for international terrorism threats? That presents another level of information. This overall view may give you immediate insight into the audience being reached by your target video stream.



Figure 4.06: A map of the viewer locations of a Facebook Live video.

Finally, after the video stream has ended, you should download a copy of the evidence. While video retrieval is explained in more detail later, here is the brief version. Navigate to the address of the metadata for your target's live stream. In our example, it would be the following URL. Replace "445432539206610" with the video ID number of your target stream.

https://www.facebook.com/video/video_data/?video_id=445432539206610

Within this text-only output, you should see a link similar to the following after a descriptor of sd_src (standard definition) or hd_src (high definition). Highlight this URL, copy and paste it into a new tab, and either right-click to download or click on File > Save page as in your browser.

https:\/\/scontent-lax3-1.xx.fbcdn.net\/v\/t42.90-29\/10000_3193194192_1447664_n.mp4

IntelTechniques Facebook Search Tool (inteltechniques.com/osint/facebook.html)

You may now be wondering how you are going to implement all of these searches in an easy format. I had the same thought and developed my own web tools to handle this task. Navigate to the above website in order to access an all-in-one option. This is the same page that was referenced earlier while explaining advanced profile search options. The entire left side of this page will allow you to conduct all of the Facebook Graph searches that were mentioned in this section. The current state of this tool is visible in Figures 4.07 and 4.08, split due to size.

The first group of searches allow you to attempt to locate a profile by email address or cellular number. An alternative way to search by cell will be explained in a moment. The next option displays the profile ID number (user number) for any individual or business profile when providing the user name. Immediately below that is a field to enter the target user number and submit in order to populate this number within the remaining search fields. This is useful when executing all possible searches on a single target.

The following section displays the individual "liked", "tagged", "event", and other information that was previously discussed. Make sure you are logged into to a Facebook account for this to function. Regardless of your browser, you may need to allow pop-ups for this to work. The next group will conduct the friends in common search that was previously explained. It will work best with user names instead of Facebook user numbers. The final group will replicate the most popular commonality searches that were discussed previously, and user numbers are required. Every field in this tool specifies the type of data needed.

The entire right half of this search page is focused toward Facebook searches outside of an individual profile. It allows you to search by name, location, employer, and other filters mentioned previously. It identifies the type of search within each search box. As an example, typing "Microsoft" in the third search option would immediately display any profiles that previously announced employment at Microsoft. The lower options allow you to combine searches. Entering "OSINT" and "Microsoft" in the "People who like... and worked at..." option would display profiles of former Microsoft employees that clicked "Like" on the OSINT Facebook page. Overall, this tool is not doing anything that you could not do on your own using the methods mentioned previously. Its purpose is to make searching easier.

The Multiple Variables option near the bottom will allow you to select multiple filters and launch the appropriate search. You can choose as many or as few of the following options to generate your own custom search. Clicking "AND" after each input will present an additional search field.

Name
Current Employer or Title
Previous Employer or Title
Current Home Location
Previous Home Location
Language Spoken
School Affiliation
Places Visited
Pages Liked
Year Born

The gender search option allows you to replicate the most popular user searches and filter by gender. It should be noted that only male and female are included here. While some users have chosen a different gender, such as trans-gender, these are a rarity. Also, be aware that some users do not choose a gender at all. Finally, the page search section is a collection of miscellaneous tools that can help with various search methods presented previously. The posts searches, employee filters, and event queries can all be found here. Notice that some of these require a user ID number while some do not, and all are appropriately marked.

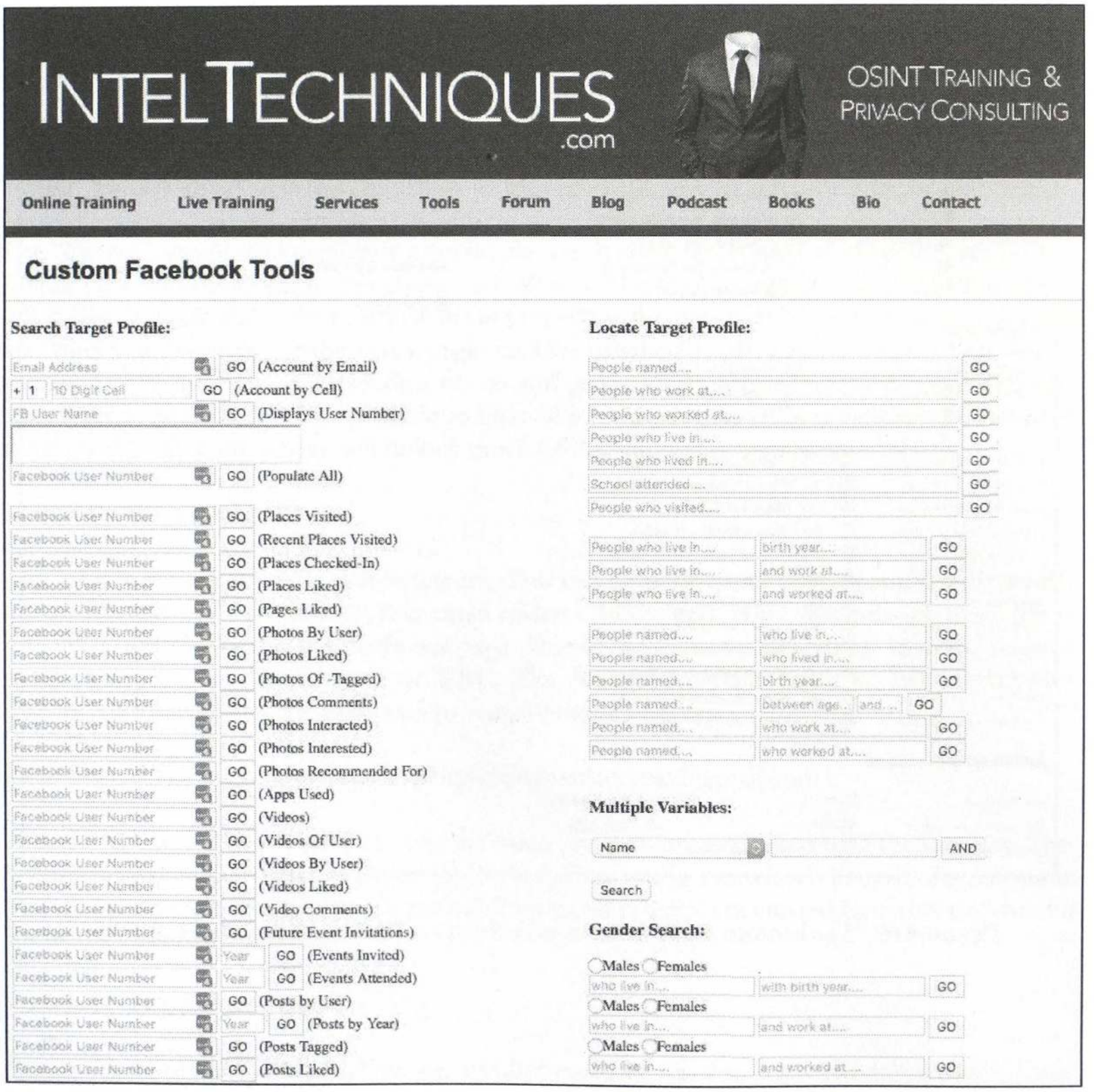


Figure 4.07: The top half of the IntelTechniques Facebook Search Tools page.

Facebook User Number GO (Employers)
Facebook User Number GO (Groups)
Facebook User Number GO (Co-Workers)
Facebook User Number GO (Friends)
Facebook User Number GO (Followers)
Facebook User Number GO (Relatives)
Facebook User Number GO (Friends' Likes)

Multiple Target Profiles:

User Number User Number GO (Common Details)
User Number User Number GO (Common Friends)
User Number User Number GO (Common Places)
User Number User Number GO (Common Check-Ins)
User Number User Number GO (Common Likes)
User Number User Number GO (Common Photo Tags)
User Number User Number GO (Common Photo Likes)
User Number User Number GO (Common Photo Comments)
User Number User Number GO (Common Video Tags)
User Number User Number GO (Common Video Likes)
User Number User Number GO (Common Video Comments)
User Number User Number GO (Common Events)
User Number User Number GO (Common Post Comments)
User Number User Number GO (Common Groups)

Friends' Info:

User Number Page ID GO (Friends by Location)
User Number Page ID GO (Friends by Likes)
User Number Page ID GO (Friends by Employee)

Facebook User Number GO (Photos Liked by Friends)
Facebook User Number GO (Photos Comments by Friends)
Facebook User Number GO (Photos of Friends)
Facebook User Number GO (Videos liked by Friends)
Facebook User Number GO (Videos Comments by Friends)
Facebook User Number GO (Videos of Friends)
Facebook User Number GO (Events Attended by Friends)
Facebook User Number GO (Apps Used by Friends)
Facebook User Number GO (Posts by Friends)
Facebook User Number GO (Posts Comments by Friends)
Facebook User Number GO (Posts Tagged with Friends)
Facebook User Number GO (Check-Ins by Friends)

Additional Information:

User Number Keywords GO (Keyword By)
Facebook User Name API Key GO (Pipl API)
Facebook User Name GO (Wayback Archives)

Detailed Search:

Posts (Keyword) GO
Posts (Keyword) GO
Photos (Keyword) GO
Videos (Keyword) GO
External Videos (Keyword) GO
Event (Keyword) GO
People that visited (Page ID Number) GO
People that checked in to (Page ID Number) GO
People that like (Page ID Number) GO
Employees of (Page ID Number) GO
Photos from (Page ID Number) GO
Videos from (Page ID Number) GO
Posts from (Page ID Number) GO
Group Members (GroupID) GO
Viewer Locations (Video ID) GO
Video Download (Video ID) GO
Video API (Video ID) GO

Video Data: ALL / Level 0 / Level 1 / Level 2

Additional Options:

Net Boot Camp

Figure 4.08: The bottom half of the IntelTechniques Facebook Search Tools page.

Once you have located the profile of your target, it should be fairly easy to navigate. If the information is public, you can click to see the friends, photos, posts, and basic information about the subject. If the profile is private, you may be limited to what you can see on this screen. Clicking the "Friends" link in the profile will load a list of the friends in the center column. This can be beneficial to identify associates of the subject. While some people will only list others with whom they have had some type of relationship, many people will list hundreds of friends that they may or may not actually know. Fortunately, Facebook gives us a search bar that we can use to filter this list. Typing a name or partial name will immediately filter the list. This works on first name or last.

When you locate a comment on a profile, never assume that the comment was not manipulated. Facebook allows a user to change the text posted at any time. Fortunately, the word "Edited" will appear under a comment if any changes have been made. Clicking this link will load a new pop-up window that will display any edits to the comment and the date and time of each edit.

The "Photos" option will display any public photos that the target has in his or her profile. The default view will list a photo identifying each album. The total number of photos in the album will be listed below the album name. Clicking on either the photo or the album name will open that album in icon view. If the user's page has been marked as private, the information will be limited. This will usually still include a photo and general wall information such as friends that have recently been added. It is possible to hide all information, but that is very rare. The search methods explained previously will unlock much of this content otherwise unavailable.

Search by Email Address

Facebook has another useful search feature. You can enter an email address and it will identify any profile that was created using that email address. In the past, you could provide this address in the main search field on any Facebook page. This technique no longer works. Instead, we must submit the search as an address, or URL. The following website address will identify any Facebook profiles that were registered to tom.smith@gmail.com.

https://www.facebook.com/search/people?q=emailaddress@gmail.com

This will link to the profile and provide the basic information associated with the account. This is also beneficial when trying to locate someone that may have a nickname instead of a real name on their profile. Additionally, it may provide alias profiles that were created by someone who did not use an alias email address.

Search by Telephone Number

Prior to November of 2012, a telephone number could be typed into the Facebook search form and the results would identify any profiles created with that number. Since Facebook usually requires a cellular telephone number to complete the signup process, this type of search yielded

very successful results. Facebook has a great database of cellular telephone numbers and associated users. Unfortunately, this feature was eliminated by Facebook after a security researcher made the vulnerability public. This feature reappeared for several months in 2014 but was later removed. It has come and gone over the past two years. Fortunately, there are a few ways around this obstacle.

Search via Facebook: Within any Facebook page, type the full telephone number, without any hyphens or parenthesis, and press enter. Often, this will display the Facebook profile associated with the target number.

Search via URL: If this does not produce a result, navigate to the following URL, replacing 6185551212 with your target telephone number.

https://www.facebook.com/search/people/?q=6185551212

Note that this example used a traditional United States number. Even then, the results can be finicky. If this fails, try the same search, but include the country code within the URL. The following would apply the U.S. country code of 1. The "+" is absolutely required when including a country code. Note that the Facebook tools mentioned previously will aid in formulating these addresses.

https://www.facebook.com/search/people/?q=+16185551212

Search via Messenger: Some readers have reported that Facebook has blocked the previous two techniques from working with some accounts. There is no explanation, but I have also witnessed this during a live course. One attendee could log into one covert account and search numbers fine, while another account on the same computer failed. The solution was to use Facebook's Messenger service at messenger.com. This chat site requires login to any Facebook account, and then allows opening of a new chat (Messenger) window. In the "To" field, you can type any telephone number and receive the associated profile information immediately. If you will be searching many numbers in a short period of time, this is the preferred method. The data appears before striking the enter key, and can be faster than the previous options.

Search via Password Reset: I urge great caution when conducting the following technique. In early 2017, a few members of my online forum began to notice that the target account was notified when executing this strategy. This could be devastating to an investigation. I have been able to reproduce this anomaly on a recent investigation. Only proceed with the following option if you are comfortable with the target being notified that someone tried to reset his account. Overall, I am surprised that the following still works. There is a Facebook website that allows you to reset your password if you have forgotten it. You must be completely logged out of Facebook before visiting this page, which can be located at the following address.

https://www.facebook.com/login/identify?ctx=recover

This will present a single search field that will accept an email address, a telephone number, a user name, or a real name. If you enter any of this information about your target, you should receive a result identifying the profile with the full name of the subject. This essentially provides the world's greatest cellular telephone number search engine. The information provided may identify other partial email addresses and will confirm the last two digits of all telephone numbers attached to the account. Do not click "Continue" on this screen, as it will send a password reset request to that individual. As a final reminder, this notification could happen even without clicking this button. This would not lock him out of the account or gain you access, but it will raise suspicion. The following summary will explain the result for each type of data supplied.

Full Name: Facebook provides real name, profile photo, partial email addresses, and redacted telephone numbers.

Email Address: Facebook provides real name, profile photo, confirmed email address, partial additional email addresses, and redacted telephone numbers.

Cell Number: Facebook provides real name, profile photo, partial email addresses, confirmed telephone number, and redacted additional telephone numbers.

User Name: Facebook provides real name, profile photo, partial email address, and redacted cellular telephone number.

The profile photos can help identify the appropriate target when searching directly on Facebook with the real name listed on the search result. This tool can also be used to identify a cellular number from a target Facebook profile in some scenarios. Consider that your target is facebook.com/jarvists. Navigating to this password recovery page, and providing jarvists as the user name presents information about the target including a real name, partial email address, and the last two digits of his cellular telephone number. The remaining numbers are redacted. If this same Facebook user was also a Snapchat user prior to 2015, there may be a vulnerability that can be exploited. For many years, Snapchat allowed users to query by cell phone number to identify potential friends. Programmers quickly learned that they could submit every possible cell phone number and collect a huge database of users. The data was leaked online and can be searched at findmysnap.com. A search of the target user name here reveals the cell number attached to the account, with the final two numbers masked for privacy. Between the Facebook password reset tool and the Snapchat leaked data test page, you can identify a complete cell phone number for millions of people.

Search via Yahoo: I rarely rely on this technique, as I do not believe it is necessary, but I want to keep it in this work as it does still function. If the previous attempts break, we may one day be required to return to the old way of doing things. A Facebook account and Yahoo email account will be required for this technique. Through a web browser, log into an alias Facebook account in one tab and log into an alias email account on Yahoo Mail in another tab. On the "Contacts" page of your Yahoo account, select "Add a New Contact". This will present a form that will allow

you to enter a target telephone number. The "Name" fields can contain any data, and I usually use numbers beginning at 001. Save this entry and go to your Facebook page. Click on the "Find Friends" button next to your name in the upper right portion of the page. This will present several options for importing friends into Facebook. Choose the last option and provide your Yahoo email address. You will need to confirm that you want to give Facebook access to your Yahoo contacts by clicking "I agree" on the pop-up window. Facebook will then identify any user profiles that are associated with the target telephone number. Facebook will encourage you to add these friends, but do not choose that option. If you do, your target will be sent a friend request from your account.

This should be a reminder about how important it is to only log into profiles that do not contain any personal information. While the target should never know that you were searching their information, an anonymous profile ensures that your identity is not compromised in case a mistake is made during the search techniques. This technique usually identifies the current users of cellular telephone numbers regardless of how the cellular number was registered. This is often successful in identifying owners of cash telephones that could not be identified through traditional resources.

Identify User's Email Address

Yahoo can provide additional information that traditional searches cannot. Yahoo can extract the email address of some Facebook users based on the address associated with the Facebook account. This method will only work if you are friends on Facebook with the target profile. Often a simple friend request to the target will result in an accepted connection. While logged into your Yahoo email account, click on the "Contacts" tab and then the "Import Contacts" button. This will present four options, and you should choose the first option titled "Facebook". You will be asked to confirm that you want Yahoo to access your Facebook contacts. Yahoo will then identify your Facebook friends and extract their information into your Yahoo contacts. This will often include the user's full name and any email addresses associated with the target.

Changed User Name

If you have an ongoing investigation, you should make note of your target's Facebook user ID number. If he or she changes the real name or user name of the Facebook profile, this user ID will always connect to the live account. A previous Graph example displayed a Facebook user number of 651620441. You can now always navigate to the following address to view this profile.

facebook.com/651620441

If this user changes his user name to bart.lorang.2 or his real name to John Doe, the above address will still find the profile. A standard search would not. This will always be the easiest way to follow the activity of a target. If you are going to create a bookmark of this subject's page, make sure it is using the number instead of the name. It is very popular with young people and online

criminals to change this information often. It introduces confusion for the authorities that are monitoring. The target's new user name updates on their friends' pages automatically.

Embedded Photos

As with any other online evidence, you should download and archive any Facebook photos of interest of your target. After you locate a photo using any of the methods explained in this chapter, click on it to expand the image inside Facebook's default image viewer. Facebook compresses all photos that are uploaded, and this view is smaller than the original uncompressed version. The best way to manually save the photo will vary based on the web browser that you are using. Overall, right-clicking and selecting "Save image as" will work.

Tracking Photos

The first two editions of this book discussed a website that would track a photo hosted on Facebook. If you encountered an image that had been posted to a non-Facebook site, but the original link was obviously from a Facebook server, you could use this service to identify the Facebook profile connected with the image. This could also be used to search links to Facebook photos that were sent via text message or chat room. This service, as well as others like it, stopped functioning in 2013. However, we can still perform this type of search manually. The following scenario should explain the process.

Imagine that you have located a photo that was posted to an online forum or blog. It could have also been sent as a text message. The link to the actual image is the following address. The following link connects directly to a photo.

https://fbcdn-sphotos-c-a.akamaihd.net/hphotos-ak-xpa1/t31.0-8/1614393_10101869091776891_1149281347468701704_o.jpg

The domain "fbcdn.net" indicates that this image is stored on a Facebook server, and that it is connected to a Facebook profile. However, we do not know the associated profile. There are three groups of numbers listed in this link, and each is separated with an underscore (_). We are interested in the second number (10101869091776891). We should construct a new address based on the following structure.

https://www.facebook.com/photo.php?fbid=NUMBER

If we enter that second number at the end of this address, it will connect us to the original photo page that possesses the image of interest. This will identify the person that possesses this picture on their profile, any tagged information, the date it was added, and any comments on the photo. The actual address for this example is the following.

https://www.facebook.com/photo.php?fbid=10101869091776891

We now know that Mark Zuckerberg originally posted this photo on his Facebook page. We can analyze the page to determine the date it was posted and any communication in reference to the image. If the photo that you connect to is private, you will receive a "This page is not available" error.

Facebook Friends Extraction

I was recently presented a Facebook scenario without any obvious solution. Assume that you find the Facebook profile of your target, and there is a large list of "Friends". You want to document this list, and a screenshot is not good enough. You want to capture the hyperlinks to each account and have a data set that can be manipulated or imported elsewhere. There are several outdated online tools that claim to "scrape" this data, but none of them work. In 2016, Facebook changed their Terms of Service (TOS) which now blocks any type of scraping of Friend data. Furthermore, the Facebook API no longer allows the display of this data either. There is a solution, and it will involve minimal geek work.

First, identify the page of your target. For this example, I will use the following public profile:

https://www.facebook.com/darya.pino/friends

She has several friends, so I will hold down the Space Bar on my keyboard to load the entire page. I will now highlight her entire friends list and use ctrl-c or right-click>copy to copy the data. I find it best to click directly above the left side of the first friend and hold until the lower right area of the last friend. The friends list should highlight. Now, open Microsoft Excel. Click on the "B" in column B to highlight the entire column. Paste the content with either ctrl-v or right-click>paste. This will appear disorganized, but the data is there. The images will be on top of the user data, which will not work for a final report.

Use F5 to launch the "Go To" menu and select "Special" in the lower left. Select "Objects" and click OK. This will select all of those images. Hit the delete key to remove them. You will now see only the text data (with hyperlinks). Now click on the "A" in column A and paste the friend content again with either ctrl-v or right-click>paste. Right-click any cell in this column and choose "Clear Contents". This will remove any text, but keep the images.

Place your mouse in between columns A and B and resize column A to be a bit larger than one of the images. Do the same with Column B to fit in all of the text. Use the "Find and Replace" feature to find every instance of "Add Friend" and replace it with nothing. This will remove those unnecessary entries. In the "Home" menu, choose "Format" and then "Auto Fit Row Height". This will eliminate unnecessary spacing. Select Column B and Left Justify the text. Your final result will be a clean spreadsheet with all of the images, names, and active links from your target's Facebook page. This is not the cleanest way of doing things, but it will work.

Facebook continuously makes both minor and major changes to their search functions. Some of these instructions may not work one day and work fine the next. Hopefully, this section has given you new ideas on ways to completely analyze your next target on Facebook.

Suspended Covert Accounts

In early 2017, I logged into a covert Facebook account that I had been reserving for a new investigation, only to discover that it was "locked". Facebook had detected unusual activity, and demanded that I provide government issued identification in the name of my covert account if I wanted it unlocked. I logged out, cleared my cache, and tried another account. Blocked again. I discovered that 20% of my accounts had been suspended due to this unusual activity. What was so unusual? Likely that I had never used them. I had created them, shelved them, and let them sit dormant for months. That does not present the appearance of a real user. I have since heard from many readers that have had the same issue. The following technique should help drastically with this problem. If you have a covert account created and working, you might consider protecting it with the following actions.

While logged into your covert Facebook account, navigate to If This Then That (ifttt.com) and create a new free account. Alias name and information should be acceptable. Navigate to the applet creation page at ifttt.com/create, and click on the big "this" option in the statement "If this then that". Click on the "RSS Feed" option and then "New Feed Item". Provide a URL from a popular online blog that posts updates frequently. For this example, we will use the RSS feed for the blog at krebsonsecurity.com/feed. I chose this account because it openly shares an RSS feed option. After providing this URL, click the "Create Trigger" option. Next, click the large "That" option and select Facebook. Click the "Create a Link Post" option and then "Create Action". Click "Finish", and you are done. Note that you may be asked to approve that you wish for IFTTT to be able to communicate with your Facebook profile during this process.

You now have an applet at IFTTT that will automatically post to your account every time that Brian Krebs publishes a blog article. In other words, every few days it will appear that you logged into your Facebook account, created a new post, and announced a link to Brian's website. This is all done behind the scenes. In my experience, this eliminates the "red flag" from Facebook when an account is no longer being used (or was never used). I now have these set up on every covert account that I possess. Since my adoption of this technique in May of 2017, I have not had a single profile become suspended.

Facebook search techniques will continue to grow, change, and be eliminated. Every month, I change something within the Facebook Search Tool due to a modification in the way that they store data. The previous search options in this chapter should provide the knowledge necessary to adapt with these changes.

Chapter Five
Social Networks: Twitter

Twitter is a social network and micro blogging service that limits most posts to 140 characters. In 2017, Twitter reported that there were 500 million Twitter posts, or "Tweets", posted every day. Basically, users create a profile and post Tweets announcing their thoughts on a topic, current location, plans for the evening, or maybe a link to something that they feel is important. A user can "follow" other users and constantly see what others are posting. Likewise, a user's "followers" can constantly see what that user is up to. The premise is simply sharing small details of your life for all of your friends to see, as well as the rest of the world. Most users utilize the service through a cellular phone, which can capture the user's location and broadcast the information if the location feature is enabled. Obtaining information from Twitter can be conducted through various procedures.

Twitter Search (twitter.com/search)

This is the official site's search interface, but it is nothing different than the search field at the top of any Twitter profile or search result. I only present this because navigating to twitter.com usually offers a signup page, but no option to search.

Twitter Advanced Search (twitter.com/search-advanced)

This page will allow for the search of specific people, keywords, and locations. The problem here is that the search of a topic is often limited to the previous seven to ten days. Individual profiles should display Tweets as far back as you are willing to scroll through. This can be a good place to search for recent data, but complete archives of a topic will not be displayed. The following explains each section.

All of these words: The order of wording is ignored here, and only the inclusion of each of the words entered is enforced.

This exact phrase: Every Twitter search takes advantage of quotes to identify exact word placement. Optionally, you can conduct the search here to get precise results without quotes.

Any of these words: You can provide multiple unique terms here, and Twitter will supply results that include any of them. This search alone is usually too generic.

None of these words: This box will filter out any posts that include the chosen word or words.

These Hashtags: This option will locate specific posts that mention a topic as defined by a Twitter hashtag. This is a single word preceded by a pound sign (#) that identifies a topic of

interest. This allows users to follow certain topics without knowing user names of the user submitting the messages.

From these accounts: This section allows you to search for Tweets from a specific user. This can also be accomplished by typing the user name into the address bar after the Twitter domain, such as twitter.com/JohnDoe92. This will display the user's profile including recent Tweets.

To these accounts: This field allows you to enter a specific Twitter user name. The results will only include Tweets that were sent to the attention of the user. This can help identify associates of the target and information intended for the target to read.

Mentioning these accounts: While these messages might not be in response to a specific user, the target was mentioned. This is usually in the form of using "@". Anyone mentioning me within a Tweet may start it with @inteltechniques.

Near this place: This field allows for the input of a zip code or city name. The default 15 miles setting would produce Tweets posted from within 15 miles of the perimeter of the zip code supplied. In a moment, I will explain a more effective search technique for location.

From this date: The final option allows you to limit a search to a specific date range. We will do this manually in just a moment.

Overall, I do not ever use the Twitter Advanced Search page. In the following pages, we are going to replicate these searches within our own custom options, which will be much more powerful than these standard solutions. The results of any of these searches can provide surprisingly personal information about a target, or generic content that includes too much data to be useful. This data can be used for many types of investigations. Law enforcement may use this data to verify or disprove an alibi of a suspect. When a suspect states in an interview that he was in Chicago the entire weekend, but his Twitter feed displays his Tweet about a restaurant in St. Louis, he has some explaining to do. Private investigators may use this content as documentation of an affair or negative character. Occasionally, a citizen will contact the authorities when evidence of illegal activity is found within a person's Tweets. The possibilities are endless. First, let's find your target's Twitter profile.

Twitter Person Search (twitter.com/#!/who_to_follow)

Locating your target's Twitter profile may not be easy. Unlike Facebook, many Twitter users do not use their real name as their profile name. You need a place to search by real name. I recommend Twitter's "Who to follow" search page. Loading this page presents a single search option under the Twitter bar that can handle any real name. Scrolling down the list I can look through the photo icons and brief descriptions to identify my target. Clicking on the user name will open the user's Twitter profile with more information. While most Twitter search options do not require you to be logged into an account, this one does.

Followerwonk Bios (moz.com/followerwonk/bio)

Twitter's Who To Follow option can be great when you know the exact name that a target used when creating an account. If you are unsure of the real name, or if the target has a very common name, Followerwonk can help you identify the profile that you are seeking. This service allows you to search Twitter profiles for any keyword that may help you locate a profile of interest. You can choose the default "Profiles" search or the focused "Twitter Bios Only" option. The "More Options" under the main search box will display numerous fields including Location, Name, and Follower details. A search of "John Smith" reveals 21,156 Twitter profiles. However, a search of "John Smith" from "New York City" reveals only 125 profiles. Filtering out profiles that have not posted at least 100 Tweets reveals only 44 profiles. This is a manageable number of profiles that can be viewed to identify the target.

Twitter Directory (twitter.com/i/directory/profiles)

If you still cannot locate your target's profile, you may need to resort to the Twitter Directory. This awkward and difficult monstrosity tries to allow you to browse through the millions of Twitter profiles alphabetically. First, choose the first letter of the first name of your target. This will present a range of options. You then need to select the range in which your target would be listed, and that selection would open a new window with hundreds of name range options such as "Mike Hall – Mike Hirsch". You will need to keep using this drill-down method until you reach a list of actual profiles that meet your criteria. I do not enjoy this method, but sometimes it is all that I have. I once found my target this way after he used a misspelled version of his real name.

If searching by real name through the previous three methods does not produce your target, your best option is to focus on the potential user name. I will discuss user name searches at length in Chapter Nine (User Names). Overall, any associated user names from other networks such as Instagram, Snapchat, and YouTube, should be tried on Twitter in the format of twitter.com/user name.

Password Reset Requests (twitter.com/account/begin_password_reset)

A method was presented earlier where a cellular telephone number was provided to a Facebook password reset page and the owner of the number was revealed. While Twitter possesses a similar lookup tool, it does not completely identify the individual. This page allows input of a telephone number, email address, or user name. The result includes the last two digits of their cellular number and a redacted email address. This data can be used to verify previously collected data or to identify the email provider used by a target, such as Yahoo or Gmail. This also verifies that the provided information is actually connected to an account.

Search Operators

Similar to the way that search engines use operators as mentioned in Chapter Three, Twitter has its own set of search operators that will greatly improve your ability to effectively search Twitter. Two of the most powerful options are the "to" and "from" operators. I use these daily in order to filter results. Consider the following examples using our target of twitter.com/sultryasian. We can obviously navigate directly to her page, but it is full of promoted Tweets, reTweets, and whatever content that she wants us to see. Instead, the following search within any Twitter window will limit our results only to her outgoing Tweets from her account. Clicking the Latest option in the Twitter menu will place these in reverse-chronological order.

from:SultryAsian

This provides a better view into her thoughts. Both her Twitter profile and this results page give us insight to the messages that she is sending out, but what about the incoming content? With most traditional Twitter investigations, we tend to focus on one side of the conversation. There is often a plethora of associated messages being sent to the attention of the target that go unchecked. In order to see all of the posts being publicly sent to her, we would search the following.

to:SultryAsian

We now see all of those incoming messages that she cannot control. While she can prohibit them from being seen on her profile, she cannot block us from this search. When I have a missing person or homicide victim, I would much rather see the incoming messages versus the outgoing. We can also combine these options to create an extremely targeted query. At first glance, I did not see many incoming messages to SultryAsian from HydeNS33k. However, the following Twitter search tells a different story. It isolates only these posts.

to:SultryAsian from:HydeNS33k

Search by Location

If you are investigating an incident that occurred at a specific location and you have no known people involved, Twitter will allow you to search by GPS location alone. The Twitter Advanced Search allowed us to search by zip code, but that can be too broad. The following specific search on any Twitter page will display Tweets known to have been posted from within one kilometer of the GPS coordinates of 43.430242,-89.736459.

geocode:43.430242,-89.736459,1km

There are no spaces in this search. This will be a list without any map view. They will be in order chronologically with the most recent at top. The "1km" indicates a search radius of one kilometer.

This can be changed to 5, 10, or 25 reliably. Any other numbers tend to provide inaccurate results. You can also change "km" to "mi" to switch to miles instead of kilometers. If you want to view this search from the address bar of the browser, the following page would load the same results.

https://twitter.com/search?q=geocode:43.430242,-89.736459,1km

You can add search parameters to either of these searches if the results are overwhelming. The following search would only display Tweets posted at the listed GPS coordinates that also mention the term "fight". Notice that the only space in the above search is between "km" and "fight".

geocode:43.430242,-89.736459,1km "fight"

It would be inappropriate to finish this section without a discussion about the lack of geo-enabled Tweets. Several years prior, this search would have been highly productive, as an alarming number of Twitter users were unknowingly sharing their locations with every post. Today, it is the opposite. The default option for Twitter is NOT to share location. A user must enable this option in order to appear within these search results. In my experience, catching a criminal from a location-enabled Tweet is extremely rare. However, we should be aware of the possibility.

Mandatory and Optional Search Terms

You may have a scenario that requires a specific search of both mandatory and optional terms. Twitter does not provide a published solution for this. However, it does support this type of search. Assume that you are investigating threats against your target named Michael Parker. You believe that people may be tweeting about him with reference to violence. Searching his name alone produces too many results. Since you only want posts that include violent terms, the following search on any Twitter page may be appropriate.

"Michael Parker" kill OR stab OR fight OR beat OR punch OR death OR die

The name within quotes forces Twitter to only give you results on those exact terms. That is your mandatory portion. The words kill, stab, fight, beat, punch, death, and die are all optional because the term "OR" is between each. This term must be uppercase, and will only require one of the optional words be present within the search results.

Date Range Search

If you are searching vague terms, you may want to filter by date. This option is now available on the advanced search page, but I believe it is important to understand how Twitter performs this task. Assume that you are investigating a bomb threat that occurred several weeks or months ago. A search on Twitter of the terms "bomb threat" will likely apply only to recent posts. Instead, consider a date specific search. The following query on any Twitter page would provide any posts

that mention "bomb threat" between January 1, 2015 and January 5, 2015.

since:2015-01-01 until:2015-01-05 "bomb threat"

My favorite use of this search technique is to combine it with the "to" operator or a name search (or both). This allows you to go further back in time than standard profile and Twitter feed searches typically allow. Consider an example where Twitter user humanhacker is your target. You can visit his live Twitter page and navigate back through several thousand Tweets. However, you will reach an end before obtaining all Tweets. This could be due to Twitter restrictions or browser and computer limitations. He currently has 11.4 thousand Tweets. Even if you could make it through all of his Tweets, you are not seeing posts where he is mentioned or messages sent publicly to him. I recommend splitting this search by year and including mentions and messages directed toward him. The following search within Twitter displays all Tweets from the Twitter name humanhacker between January 1, 2012 and December 31, 2012.

from:humanhacker since:2012-01-01 until:2012-12-31

This may create a more digestible collection of Tweets that can be collected and archived appropriately. There may be no other way of identifying these messages since you cannot likely scroll back that far. In my investigations involving targets with several thousand posts, I conduct multiple searches within Twitter that span several years. The following would collect yearly sets of Tweets posted by humanhacker since 2006.

from:humanhacker since:2006-01-01 until:2006-12-31
from:humanhacker since:2007-01-01 until:2007-12-31
from:humanhacker since:2008-01-01 until:2008-12-31
from:humanhacker since:2009-01-01 until:2009-12-31
from:humanhacker since:2010-01-01 until:2010-12-31
from:humanhacker since:2011-01-01 until:2011-12-31
from:humanhacker since:2012-01-01 until:2012-12-31
from:humanhacker since:2013-01-01 until:2013-12-31
from:humanhacker since:2014-01-01 until:2014-12-31
from:humanhacker since:2015-01-01 until:2015-12-31
from:humanhacker since:2016-01-01 until:2016-12-31
from:humanhacker since:2017-01-01 until:2017-12-31

This same technique can be modified to display only incoming Tweets to humanhacker for these years. Replace "from" with "to" to obtain these results. The 2008 posts would appear as follows.

to:humanhacker since:2008-01-01 until:2008-12-31

You can combine all of these options into a single result, but I only recommend this after you have attempted the more precise options mentioned previously. While the next search should

theoretically display all of his outgoing Tweets, incoming Tweets, and mentions, it is not always complete. The following would include this data for 2008.

"humanhacker" since:2008-01-01 until:2008-12-31

There are many uses for a date range search. Any supported Twitter search should work combined with dates. This might include a location search for a specific date related to an investigation. As a test of the possibilities, consider that you want to identify an email address for this target. His live Twitter page will not reveal this, as he no longer posts his email, likely to prevent spam. However, the following search is quite productive.

from:humanhacker email since:2006-01-01 until:2009-12-31

This isolates only his posts from the beginning of Twitter until the end of 2009. Only four results are present, including the Tweet as seen in Figure 5.01. We will use this data during our automation process as discussed in Chapter Twenty-Two (sorry Chris).

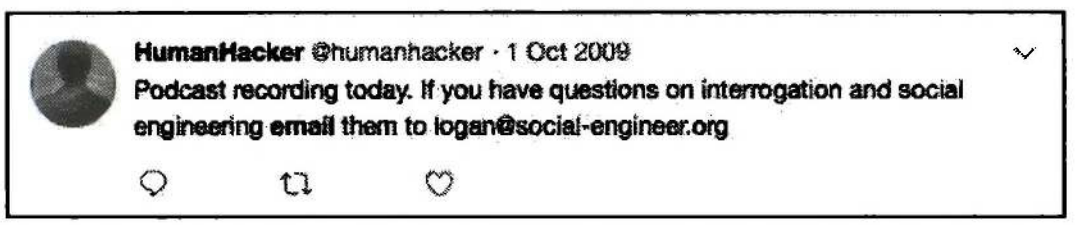


Figure 5.01: An old Twitter post including an email address of the target.

Media and Likes

You may want to filter all results from a target Twitter profile and only see those which have some type of media embedded. There is not a search operator to force this, but the following direct URL will display only these posts.

https://twitter.com/humanhacker/media/

A trend that has seen rapid adoption over the past few years is the "Liking" of posts. When a user wants to acknowledge something said by another user, but does not necessarily want to respond or reTweet, clicking the small heart icon indicates that the post was "liked". The following direct URL displays all of the posts that our target liked.

https://twitter.com/humanhacker/likes/

Deleted, Suspended, and Missing Tweets

Twitter users may delete their own accounts if there is suspicion that an investigation is under way. If this happens, searching on Twitter will not display any of the posts. Furthermore, a person

might only delete individual Twitter posts that are incriminating, but leave non-interesting posts on the profile to prevent raising suspicion associated with deleting an entire account. Some users may find their accounts suspended for violating Twitter's terms of service. In any of these scenarios, it is still possible to retrieve some missing posts using various techniques.

If I encounter a Twitter user that has recently deleted some or all of their messages, I conduct a cache search of their profile. There are various ways to do this, and I will demonstrate the most common. In this example, I conducted a search on Twitter for "deleted all my Tweets" on December 15, 2017. This provided many users that recently posted that they had just deleted all of their content. This helped me identify a good target for this type of demonstration. The first user I located was "WestCornfield". He had one Tweet and it referenced deleting all of his posts, and it is seen in Figure 5.02.

I attempted a search on Twitter of from:WestCornfield, which provided no results. I conducted a search of to:WestCornfield, which provided dozens of incoming messages from his friends. This was a good start. I then went to Google and conducted a search for "Twitter WestCornfield". The first search result was a link to the user's live Twitter page. Instead of clicking on the link, I chose the Google Cache view of his profile by clicking the small green "down arrow" next to the URL and selecting "Cached". This view identified twenty deleted Tweets from this account. Two of these posts can be seen in Figure 5.03. Google identified this capture as taken on December 12, 2017.

This may be enough for your investigation. Occasionally, I need to identify content that was deleted weeks or months before my investigation. The previous technique will likely not provide much assistance because the Google Cache is probably a recent copy of their live page. The cache may be missing Tweets you want to see. I next replicated this process on Bing and Yandex. Bing's cached view was taken on December 7, 2017 while Yandex's cached view was collected on November 3, 2017. Each of these possessed unique posts and images. Figure 5.04 displays a recovered post from Bing. Next, I returned to Google to obtain further data. I searched the following, which provided only results that possess a URL that begins with twitter.com, then my target's user name, then "status". This will force Google to present direct links to actual Twitter posts.

site:twitter.com/westcornfield/status

The result was 56 posts. When I clicked on each of these, Twitter informed me that the post had been deleted. However, opening the cached version from the Google result displayed each of the posts. In Figure 5.05, you can see that Google is now identifying deleted posts as far back as October 2017. This process should also be repeated using the cached view options of Bing and Yandex. Next, we should check the Wayback Machine as mentioned in Chapter Three. If you recall, you can search their archives by keywords or direct URL. The following address connects us directly to their archive of his account.

http://web.archive.org/web/*/twitter.com/WestCornfield

This identified a capture of his profile on December 6, 2017. Opening this archive displayed his Twitter profile dating back to November 8, 2017. Figure 5.06 displays this deleted Tweet.

While our target removed his content from his profile, he did not remove his history. In order to see the Twitter posts that he had previously liked before wiping out his page, we can navigate to the following URL. In this example, we see the hundreds of messages that identify his interests.

https://twitter.com/WestCornfield/likes/

While every investigation is unique, I wanted to demonstrate the importance of checking every source. These searches took less than three minutes using my custom Twitter search tool discussed in the next section. While you will likely never rebuild an entire deleted account, the posts obtained with this technique are something you did not have before.

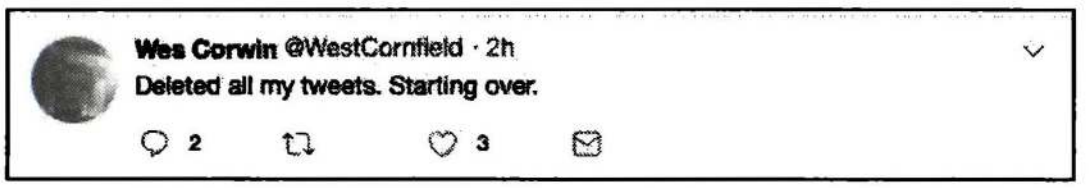


Figure 5.02: A live Twitter post announcing Tweet deletion.

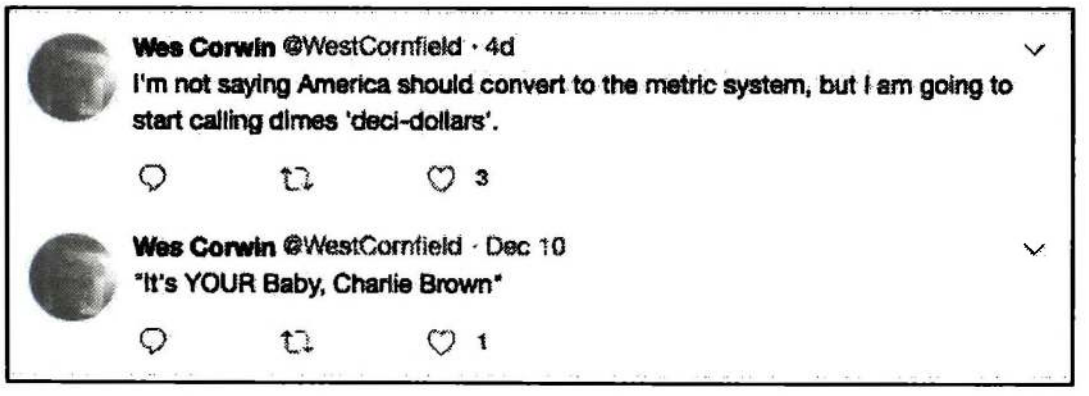


Figure 5.03: Google cached Twitter posts recovered after deletion.

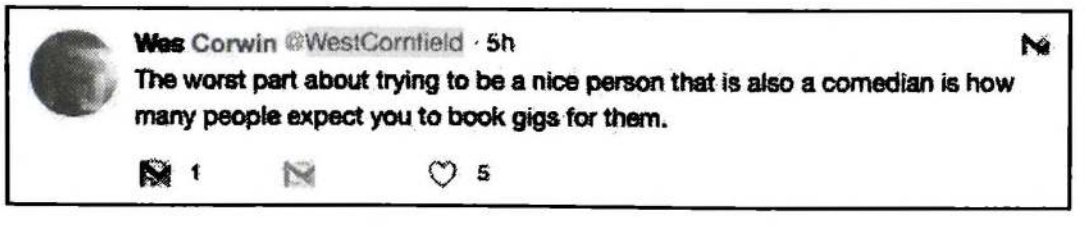


Figure 5.04: A Bing cached Twitter post recovered after deletion.

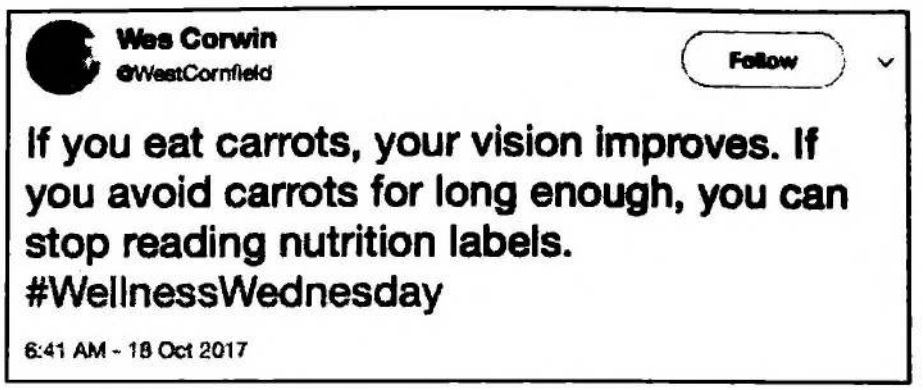


Figure 5.05: A Google cached Twitter message URL of a deleted post.

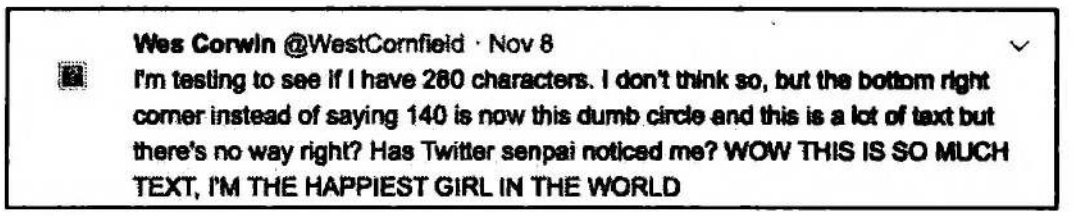


Figure 5.06: A recovered deleted Tweet from the Wayback Machine.

Twitter Bio Changes (spoonbill.io)

Similar to the way that users delete Tweets and comments, they also modify the information within their Twitter Bio on their profile page. Several sites have come and gone which attempt to record these modifications, and my current favorite is Spoon Bill. Use of this free service from the home page requires you to log into your Twitter account. However, a direct URL query will display any stored results. If you were researching my Twitter handle, the following address would bypass the account login.

https://spoonbill.io/data/inteltechniques/

This page displays several changes I made to my account on 9/8/16, including changing my name and location. Searching more active people will usually reveal many pages of changes, and will always display the previous content as stricken, and highlight any changes in green.

Real World Application: In 2017, I assisted a law enforcement agency with a missing juvenile case. Authorities had suspicion that she may have initially ran away, but became more concerned when all of her social networks became dormant and her cell phone stopped being used. Her Twitter profile possessed no valuable Tweets, and I could not find any deleted content to recover. I supplied her name to Spoon Bill and immediately received a log of changes to her Twitter profile. Two months prior to her disappearance, she listed her Snapchat user name on her bio. This led me to her Snapchat profile photo. This photo and user name had not been present in any of my previous investigation, and revealed a whole new world of leads. The user name was associated with an email address, which was used by the missing person to create another Facebook profile. The user name from Facebook revealed her user ID number, which was used

in the Facebook search tools mentioned previously. The "hidden" images connected to this account provided many interesting suspects. Within an hour, a single suspect had emerged from the new discoveries, and the juvenile was located safely nearby. I cannot overstate the necessity to retrieve modified and deleted content during every Twitter investigation.

First Tweet – Keyword (http://ctrlq.org/first)

If you are ever following a trending hashtag or unique keyword, you may want to know the user that first posted the topic. This is not available natively within Twitter, but you can use this website to perform the search. A result for the first time a person used the hashtag of OSINT (#OSINT) reveals the message and the date of October 27, 2008. I once used this to determine the person responsible for posting an inappropriate photo of a minor that included very specific text within the post. It had been reTweeted many times, and this tool quickly locate the first post.

First Follower (socialrank.com/firstfollower)

It may be beneficial to know the first follower of your target on Twitter. This will likely be the person that introduced the target to Twitter and can often identify a former associate. If you want only this one piece of information, First Follower will usually give you the result. However, I have found that you must often click the search option several times before the result is displayed.

IntelTechniques Twitter Search Tool (inteltechniques.com/osint/twitter.html)

I found myself using many of these manual Twitter techniques daily. In order to prevent repetitive typing of the same addresses and searches, I created a custom web page with an all-in-one solution. Navigate to the website above to access this resource. This page includes embedded JavaScript that will structure and execute web addresses based on your provided information. Figure 5.07 displays the current state of this tool. Any utilities that do not look familiar will be described in the remaining pages of this book.

This tool will replicate many of the Twitter searches that you have read about here. The first option on the left side populates all of the search fields with your target's Twitter name when entered there. This should save time and avoid redundant pasting of data. Clicking "Go" next to each option executes the query in a new tab in your browser. As a reminder, this page collects absolutely no information from your searches. They are all conducted within your own browser, and no data is captured by my server.

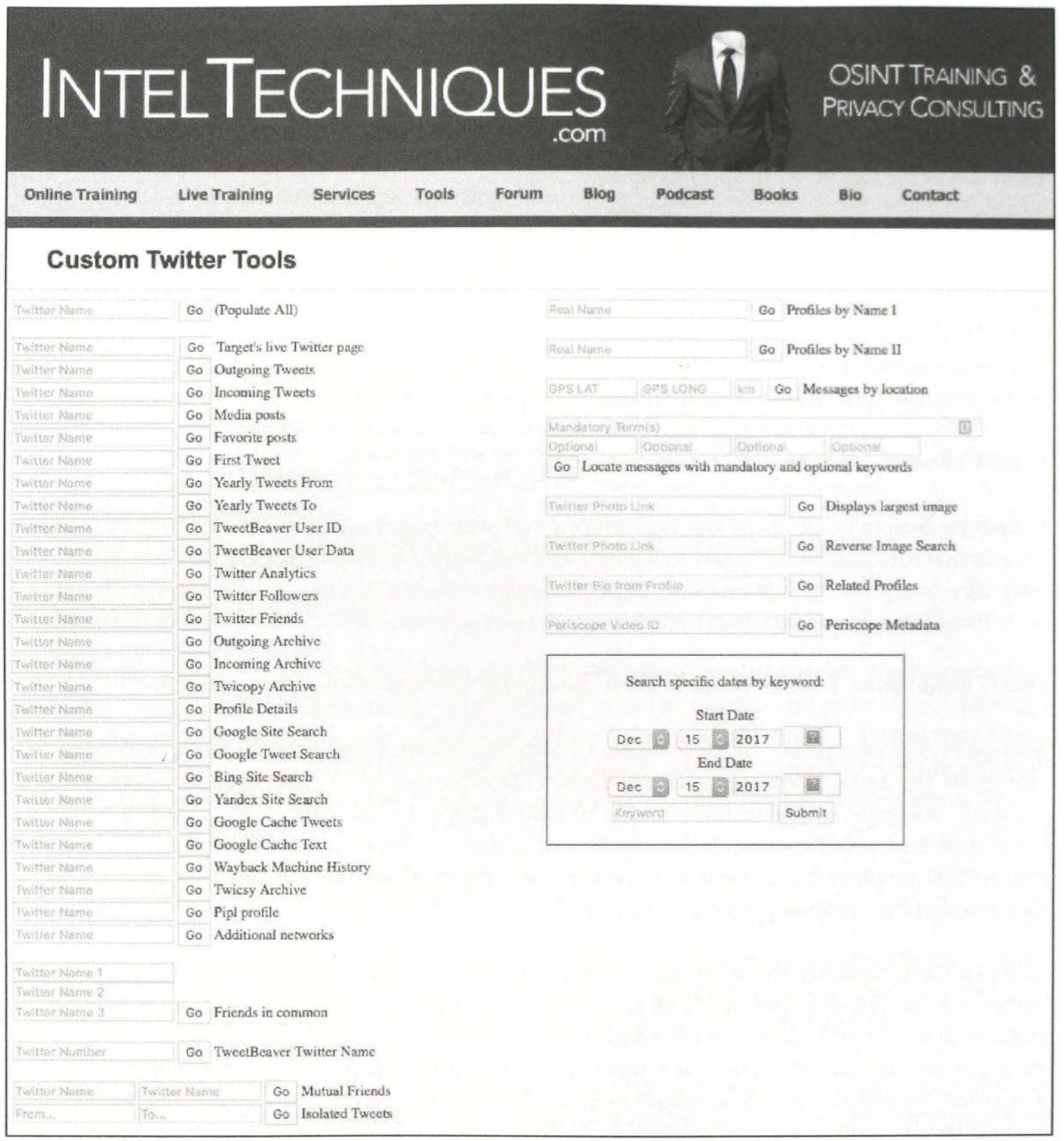


Figure 5.07: The IntelTechniques Custom Twitter Search Tool.

TweetBeaver (tweetbeaver.com)

This is currently my absolute favorite third-party Twitter reporting tool. It is the most robust option available to us for exporting content from an account or researching associations. There are currently ten unique options within this site, and I will explain each of them with usage scenarios. Note that you must be logged into a Twitter account for any of these to work, as they all leverage the Twitter API in order to function. Please only use covert accounts, and never your own personal Twitter sign-in.

Convert Name to ID: This is the most reliable way to take a Twitter user name, such as jms_dot_py, and convert it to the corresponding user number, which is 817668451. This can be vital for investigations. Users can always change their user name at any time, but the user number cannot be modified.

Convert ID to Name: This is the opposite of the above technique. If you had previously identified jms_dot_py as your target Twitter account only to find later that he had changed his user name, you could easily locate his profile by providing the assigned user number (817668451).

Check if two accounts follow each other: As the title implies, this option quickly sorts out whether two users follow each other. An actual output appears below.

@inteltechniques does not follow @jms_dot_py
@jms_dot_py follows @inteltechniques

Download a user's favorites: This is the first tool where we can choose to either display the results on the screen or download them as a CSV spreadsheet. This option simply extracts a user's favorites (or likes) as discussed earlier. The results include the original author, date and time, text of the message, a direct URL to the post, and the author's bio, as seen in Figure 5.08.

Tweet author	Date posted	Text text	URL	Tweet author's biography
@TerrorFanatics	Wed Dec 06 02:20:06 +0000 2017	https://t.co /S4MDS86jLf	www.twitter.com/TerrorFanatics/statuses/938231545235636225	Tweeting (and retweeting) the best #Horror articles, videos, memes, writers, podcasts, filmmakers, etc.

Figure 5.08: A TweetBeaver result for a user's favorites.

Search within a user's favorites: If the previous technique produces too many results, this option allows you to filter by keyword. Since you could search within the file you downloaded or on the screen of results, I find this feature to be of little use.

Download a user's timeline: This may be the most useful of all of these options. Provide a target Twitter name and TweetBeaver will extract the most recent 3,200 posts from the account. Furthermore, it will include date and time of each and the direct URL to the message. When I have a Twitter target of interest, I run this tool on the account daily. It has helped me obtain the new posts every day, and identify previous posts deleted after my initial acquisition. Figure 5.09 displays the first line of a result.

Tweet author	Date posted	Text text	URL
@jms_dot_py	Fri Dec 15 20:58:36 +0000 2017	@_prasket @TeriRadichel thanks so much you guys!	www.twitter.com/jms_dot_py /statuses /941774513942945792

Figure 5.09 A TweetBeaver user timeline result.

Search within a user's timeline: Similar to the favorites search tool, I find this one redundant.

Get a user's account data: This utility provides a great summary of the account information publicly available from any Twitter account. The benefit of this method of obtaining the data is that it is quick and presented in a standard view. I can collect this information about many users, and all results will have the same format. This can aid in presentation during prosecution. Figure 5.10 displays the actual result from this target.

Screen name	Twitter ID	Name	Biography	Account created date
jms_dot_py	817668451	Justin Seitz	Creator of @Hunchly. Blogging & training #OSINT techniques. Wrote a couple of @nostarch books. @Bellingcat contributor. @C4ads fellow.	Tue Sep 11 15:44:20 +0000 2012
Location	**URL**	**Time zone**	**Geo enabled**	**Language**
Saskatoon, Saskatchewan	http://automatingosint.com/blog	Central Time (US & Canada)	not set	en
Verified	**Tweets**	**Followers**	**Friends**	
not verified	9041	7095	2112	

Figure 5.10: A TweetBeaver user account export.

Download a user's friends list: This option collects a target's list of accounts that he or she follows. This is similar to a typical friends list on Facebook, but approvals on either end are not required.

Download a user's followers list: This is a list of the people that follow the target on Twitter. This is less likely to contain actual friends, but all associated accounts should be investigated.

Bulk Account Data Download

In early 2018, TweetBeaver introduced a bulk lookup feature which can accept up to 15,000 Twitter user names. This is quite impressive, and I have put it through many tests. In Chapter Twenty-Two, you will read about an automated solution for identifying numerous Twitter accounts of interest. The bulk feature on TweetBeaver allows you to input all of the accounts identified within one query. As a demonstration, I added every Twitter account mentioned in this chapter up to this point into the TweetBeaver bulk lookup utility. The entry appeared as follows.

inteltechniques	SultryAsian	humanhacker
jms_dot_py	HydeNS33k	WestCornfield

This immediately created a CSV spreadsheet which was downloaded to my computer. Below are screen captures of the data. This gives me an immediate view into the accounts. If I had hundreds or thousands of Twitter user names, this would allow me to sort by location or popularity. I could also sort by creation date in order to identify newly created accounts. This is possibly the most useful third-party tool when you have numerous accounts of interest.

Screen Name	Twitter ID	Name	Description	Account created date
@screen_name	ID 257644794	Michael Bazzell	Open Source Intelligence (OSINT) Training and Tools. International Privacy Consultant.	Fri Feb 25 21:46:04 +0000 2011
@screen_name	ID 817668451	Justin Seitz	Creator of @Hunchly. Blogging & training #OSINT techniques. Wrote a couple of @nos	Tue Sep 11 15:44:20 +0000 2012
@screen_name	ID 3308415728	SultryAsian	Known to most as M. Not as serious as I look.	Fri Aug 07 03:49:08 +0000 2015
@screen_name	ID 7711241370	Jek Hyde	Infosec Auntie ðŸ¤ Red Team Analyst at Fortune 1 ðŸ•µï¸ Thief ðŸ”‘ Security Quality	Wed Aug 31 23:14:55 +0000 2016
@screen_name	ID 46998400	HumanHacker	This is the official Twitter account of all things SEORG - The SEVillage, SEPodcast, and t	Sun Jun 14 00:47:39 +0000 2009
@screen_name	ID 368971022	Wes Corwin	The Dean Malenko of Stand-Up Comedy and The Kazuyuki Fujita of Roast Comedy. See	Tue Sep 06 14:58:56 +0000 2011

Location	URL	Time Zone	Geo-enabled	Language	Verified	Tweets	Followers	Following
Washington D.C.	http://inteltechniques.com	Eastern Time (US & Canada)	not set	en	verified	266	5456	0
Saskatoon, Saskatchewan	http://automatingosint.com/blog	Central Time (US & Canada)	not set	en	not verified	9163	7316	2123
not set	not set	Pacific Time (US & Canada)	not set	en	not verified	2248	1894	622
Dallas, TX	not set	not set	not set	en	not verified	3820	12817	404
USA	http://www.social-engineer.org	not set	not set	en	not verified	11498	25274	273
Dallas, TX	http://wescorwincomedy.wordpress.com	Pacific Time (US & Canada)	enabled	en	not verified	42	2725	2193

Real World Application: The day this feature was released, I was investigating a suspicious Twitter account associated with violent threats toward a celebrity. The suspect had sanitized the account which prohibited obtaining valid location data. However, the suspect had numerous followers and people he was following. Using TweetBeaver, I could extract these accounts easily and supply them to the bulk lookup utility. I then sorted his friends by location which revealed a strong presence in a specific Midwest city. Of those target accounts of interest, I could see that a few were Geo-enabled. This provided home addresses for two people. The person search tools described later identified full names and telephone numbers of these two friends. Group photos found online of these targets identified one person always nearby, but without a Twitter account associated with this group. The same photos on Facebook identified my new target by name. This process led to a positive identification of my suspect. Bulk search tools help tremendously. If you only use one third-party Twitter analysis tool, I recommend TweetBeaver.

Location Information by User

Most people post to Twitter through their smartphones. This allows users to take advantage of location aware apps such as Foursquare, which they can use to "check-in" to places and let their friends know their location. While privacy-aware individuals have disabled the location feature of their accounts, many users enjoy broadcasting their location at all times. Identifying a user's location during a Twitter post is sometimes possible through various methods. Prior to 2014, identifying the GPS details of every post of many users was simple. Today, most of these identification techniques are no longer working. Please note that the following limited methods will only work if your target has not disabled the geo-locate settings within Twitter, or they previously posted historic messages before disabling location options.

GeoSocial Footprint (geosocialfootprint.com)

This service will only search for location data within the most recent 200 posts. It is a good resource for identifying the current or recent location of a target. Using this website is fairly straightforward and does not require a Twitter account. I have found that the results are not always reliable, but worth searching. On the main page, enter the target Twitter user name and click "Retrieve Tweets". This will produce a map with markers identifying the most recent locations of that target. You can click on each marker to see the content of the message, but not the date and time. This overall view provides a quick glimpse into the general location of the target.

In order to view more detailed information about these posts, click on the "Download Tweets" button. This will generate a plain text file with no file extension titled "Download". Open this file within a text viewer, such as Notepad, to display the content. The text below is the actual content received during the previous search. The message contents were redacted for space and privacy. Note that you can now view the actual GPS coordinates, date, and time of each message. This could also be imported into a spreadsheet or database.

```
Sat Jan 10 09:22:46 +0000 2015,37.7296498,-122.4118834,Having our second unit be able to
Sat Jan 10 09:21:30 +0000 2015,37.7296699,-122.4118867,These Warrior games have been amazing.
Sat Jan 10 09:20:45 +0000 2015,37.7296699,-122.4118867,"@ItsFoodPorn: The whole point of dating
Sat Jan 10 09:18:08 +0000 2015,37.7296397,-122.4117785,You'd almost swear Imma thoughtful person
```

Tweetpaths (tweetpaths.com)

You will need to sign in to Tweetpaths using your Twitter account information in order to conduct any searches. The results will be limited to the most recent 75 Tweets. However, you can use advanced options to expand this restriction. Similar to GeoSocial Footprint, input your target's Twitter name and the resulting map will focus on the geographical area of where the target was located while posting. The pop-up window identifies detailed information about a selected message. If you want to search a specific range of dates, click the "Show advanced options" link directly below the search field. This will expand the menu and allow you to enter a

start date and end date. This can allow you to expand your search beyond the 75 Tweet limit. If your target is an active user, you may want to search one week at a time and document the results. Similar to GeoSocial Footprint, the results are not always reliable. In my experience, you will receive less than 10% of the possible location tagged results. One alternative to this service is Tweet Mapper (keitharm.me/projects/tweet), which can be used to verify any results.

A great feature of Tweetpaths is that you can identify several users on the same map. Figure 5.11 displays a map with Google's satellite view enabled. Three Twitter users were searched and the results identify each with a different color of marker. This can be beneficial by showing three or more targets that were present at the same location during an incident. A user name is required to see any activity with this service. Several new websites have surfaced that will allow us to view recent Tweets from a specified area without knowledge of any user names involved.

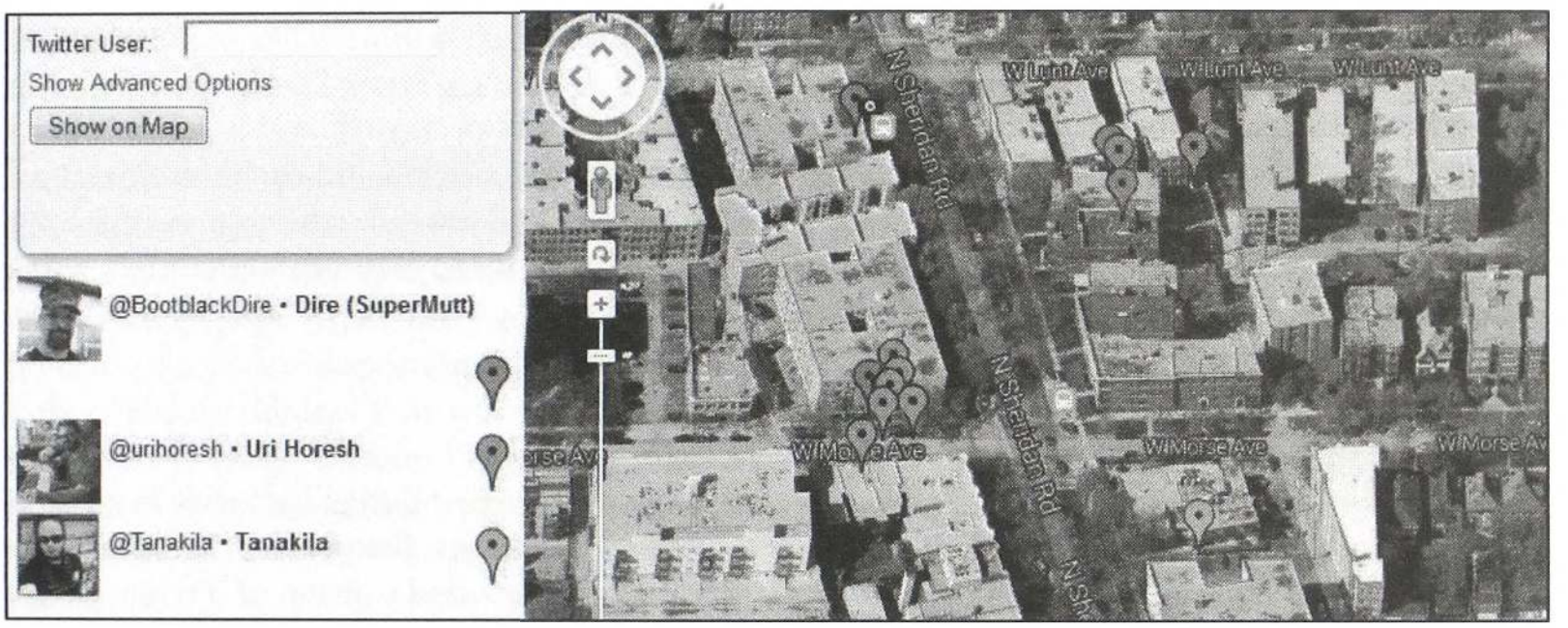


Figure 5.11: A Tweetpaths search result with satellite view and multiple accounts.

Information by Physical Location

A manual Twitter search method for identifying posts by location was explained earlier. That technique is best for recent or live information, and is limited to only recent posts. You may have a need for historical details from previous posts from a specific location. I have had better success with historic data than current content in regards to geolocation. I believe this is because most people unknowingly shared their location while Tweeting for many years. When Twitter changed the default location option to "Disabled", most users never intentionally re-enabled the feature. Fortunately, there are third-party websites that collected this historic data, and can assist with easy search. The following options will work best when you are investigating events that occurred several years prior. While you may get lucky and receive some recent posts, the majority could be quite old.

MapD: MIT (http://mapd.csail.mit.edu/tweetmap/)

MapD is a massive database platform developed through collaboration between MIT and Harvard. Currently, each college has their own interface into this data, which supplies Twitter post locations from historic Tweets. Each interface provides new ways of searching information. This website can search by topic, user name, or location. It can also combine all three options to conduct a detailed search. The first search field, titled "What", can accept any keywords and identify location-enabled Tweets that meet the criteria. The second field, "Who", can only accept a Twitter user name and will not function with a real name. The third search field, titled "Where", can accept GPS coordinates, a zip code, or a street address. This data will zoom the map to a more detailed level. The column on the left displays the Twitter message content that is represented in the map view. Finally, the graph at the bottom right identifies message volume in the chosen area. Clicking the right arrow button at the bottom will "play" the map and highlight the Tweets in the chronological order of post.

When you first load this website, a dark map is displayed without satellite view. This can be changed by clicking on the blue plus sign in the upper right area. Clicking on each visible dot will display the details of the message. This will include exact date and time, the message content, and the user that posted the Tweet. The area in the lower left is a word cloud that identifies words that appear in messages more commonly than others.

MapD: Harvard (http://worldmap.harvard.edu/tweetmap)

Harvard's version of this data interface is more basic with fewer options. This may not sound like a desirable trait. However, the clean interface makes room for a feature that may not work properly on MIT's contribution. The page displays a small box around a group of Tweets posted from a location on the map. This square is executed when clicking on any portion of a visible map. Any Twitter posts existing within the square will be presented in a pop-up window that will identify the Tweets and all available details about them. The box at the bottom will allow you to filter the results by date or keyword. The view from a monitor with decent resolution will provide a better spacious view.

Overall, both of these services are displaying identical data collected over several years. You should experiment with both until you gain your own personal preference. I believe that the MIT option is the more advanced and preferred route to go. In order to view live streaming data, we will need to use another service. I have found these services to be valuable when researching deleted accounts. If a suspect wipes out all messages on a Twitter profile, these two services may have collected the content. A search of the target's Twitter name should reveal the results.

One Million Tweet Map (onemilliontweetmap.com)

This service only displays the most recent one million Tweets on an international map. They do not have access to every Tweet available, often referred to as the "firehose", but they offer new

Tweets every second. I would never rely on this map for complete data about a location. However, monitoring a large event can provide live intelligence in an easily viewed format. I recommend using a mouse scroll wheel to zoom into your location of interest. Once you are at a level that you can see single Tweets, you can click any of them to see the content. The page will automatically refresh as new information is posted.

Search by Email Address

Technically, Twitter does not allow the search of a Twitter user by providing an email address. If you attempt this type of search on Twitter, you will receive no results, even if a profile exists for a user with that email address. To bypass this limitation, you can use a feature offered by online email providers that will communicate with Twitter.

This technique will require a Twitter account and Yahoo email account. Through a web browser, connect to an alias Twitter account in one tab and connect to an alias email account on Yahoo Mail in another tab. If you do not have a Yahoo email account, the creation process is easy at mail.yahoo.com. On the "Contacts" page of your Yahoo account, select "Add a New Contact". This will present a form that will allow you to enter the email address of the target. The "Name" fields can contain any data, and I usually use numbers beginning at 001. Save this entry and go to your Twitter page. Click on the "Find Friends" option in the settings menu. This will present several options for importing friends into Twitter. Choose the Yahoo option and provide your Yahoo email address. You will need to confirm that you want to give Twitter access to your Yahoo contacts by clicking "I agree" on the pop-up window. Twitter will then identify any user profiles that are associated with the target email address. Twitter will encourage you to add these friends, but do not choose that option. If you do, your target will be sent a notification from your account.

Embedded Photos

It is very common for users to include digital photos within their Tweets. There are multiple ways to do this. Regardless of the method, it is always important to save a copy of any photos of interest to your investigation. This section will explain the ideal process to obtaining the largest possible photos. The most likely way that people will add photos to their Twitter posts is through the native Twitter photo display. If the person used this embedded photo sharing ability, you will see a small version of the actual photo or a link within the stream of data in a person's feed. Click on this photo and it will open in a pop-up window as a larger view. You can now right-click on the photo and save the file. However, this will not be the largest size available through Twitter. Instead, right-click the photo and select "open image in new tab". This will launch a new tab that should only contain the actual image. The address of the image will likely end in a file extension such as jpg. At the very end of the URL for this image, add ":orig" without the quotes. It should open a new version of the photo. This version will likely appear larger and have a higher resolution.

As an example, I located an image at the following address from Twitter. I then added ":orig" at the end and received a much more detailed photo. The second link demonstrates this structure.

https://pbs.twimg.com/media/BstXzBYCAAECMmL.jpg
https://pbs.twimg.com/media/BstXzBYCAAECMmL.jpg:orig

There are several third-party options that allow you to extract these types of photos. I do not like to rely on them because their services do not always work properly. I prefer to archive the Twitter information from the actual source. More documentation and archiving options will be explained later.

All My Tweets (allmytweets.net)

This website provides a clean display of all of a user's Twitter posts on one screen. It will start with the most recent post and list previous posts on one line each. This view will display up to 3200 messages on one scrollable screen. This provides two analysis methods for digesting large amounts of data. Holding CTRL and F on the keyboard will present a search box. Any search terms can be entered to navigate directly to associated messages. The page can also be printed for archiving or distribution to assisting analysts. This is my preferred way of reading through the Tweets of an active user. This also prevents you from constantly loading more Tweets at the end of every page throughout a profile. While I prefer TweetBeaver for this task, it is always good to have options.

Conweets (conweets.com)

All My Tweets can be a great resource for locating posts published by your target. However, it will only show you one side of the conversation. You would need to perform the same action on every user associated with the target. Conweets may eliminate this problem. This website allows you to enter two Twitter users and it will identify the conversations between them. It will display their Tweets in order from most recent to the oldest. Each section will identify the person that started the conversation and the date. The appearance is similar to a back-and-forth text message session. You will only receive publicly visible Tweets. At the time of this writing, the Conweets website would not load, but the Twitter page was still active. This service has stopped working and returned fully functional several times. Therefore, this instruction was left here in hopes that the service will return.

Twitter Archiver (labnol.org/internet/save-twitter-hashtag-tweets/6505)

Twitter Archiver may be the simplest online tool for saving and updating Twitter streams. It is a Google add-on that takes a few minutes to setup, and the above website contains a complete video tutorial. The tool uses a Google Docs spreadsheet to capture all Tweets that match supplied search terms. You can use the tool to monitor Tweets around any hashtag, learn what people are saying about a keyword; track popular search terms; save Tweets from any geographic location;

or monitor a specific user. After you install the Twitter Archiver from this site, it will create a new Google Spreadsheet within your Google Drive account. Go to the Add-on menu, choose Twitter Archiver, and select the Authorize menu. Allow the Google Sheet to access Twitter on your behalf as the app needs this permission to fetch Tweets. It will never post anything to your Twitter account. Once your Twitter account is authorized, go to the Twitter Archiver menu again and create a new search rule. It will appear very similar to the Twitter advanced search page explained earlier. You can create rules that mention certain search terms, exact phrases, users, or any other Twitter supported search that was explained earlier.

After you have created your Twitter search query, click the "Start Tracking" button to initialize the Twitter Archiver. Internally, the sheet will connect to Twitter and pull in the historic Tweets that match your search term(s). It writes these Tweets in a separate sheet inside the Google Spreadsheet. After the initial set is pulled, the archiver will poll Twitter every hour and pull in the matching Tweets that have been posted since the last run. In addition to Tweets, the Twitter Archiver app will also import other data including the Tweet's reTweet & favorite count, the user's friend & follower count, and whether they are verified or not. If you would like to stop archiving Tweets for a particular search term, go to the Twitter Archiver menu, choose the Saved Searches menu, and you'll see a list of your existing saved searches. Select the one you wish to delete from the dropdown and hit the Delete button.

Overall, this service is extremely robust and valuable. I highly recommend watching the tutorial video to obtain a complete understanding of the many features. Your pages will automatically update while you move on to other searches. The results will be waiting for your analysis when you return. Exporting the results is easy with support for standard CSV files that can be later opened with Microsoft Excel, or imported into any other utility.

Sleeping Time (sleepingtime.org)

This site allows for a search of an exact Twitter profile name, and provides the average time period that this user sleeps. The historical Tweets are analyzed according to the times of posts. Data is then presented that suggests when the user is usually sleeping due to lack of posting during a specific time period. A query of Kevin Mitnick revealed that he is likely to be sleeping between 12am and 7am according to his Tweets. Although the idea was probably executed as a fun site, it can be quite useful.

Real World Application: Police often want the element of surprise on their side when contacting suspects. Whether this is to execute a search warrant or simply contact a subject when he or she is most likely to be home, knowing the habits of the individual can be beneficial. Locating a possible sleep pattern of the individual will decrease the chances of showing up at an empty house, only to discover that the subject works a strange night shift somewhere. Sleeping Time may have alerted an investigator that the average sleep time is 2pm to 10pm, creating an opportunity to catch the subject at home. This also works for process servers, private investigators, bill collectors, and even salesmen.

Tweet Deck (tweetdeck.com)

Tweet Deck is owned by Twitter, and it can take advantage of the Twitter "Firehose". This huge stream of data contains every public post available on Twitter. Many Twitter services do not have access to this stream, and the results are limited. Tweet Deck requires you to create and log into an account to use the service. This user account is not the same as a Twitter account. The "Create Account" button on the website will walk you through the process. Alias information is acceptable and preferred. The plus symbol (+) in the upper left area will add a new column to your view. There are several options presented, but the most common will be "Search" and "User". The "Search" option will create a column that will allow you to search for any keywords on Twitter. The following is a list of search examples and how they may benefit the investigator:

"Victim Name": A homicide investigator can monitor people mentioning a homicide victim
"School Name": A school can monitor anyone mentioning the school for suspicious activity
"Subject Name": An investigator can monitor a missing person's name for relevant information
"Event": Officials can monitor anyone discussing a special event such as a festival or concert

The "User" option will allow you to enter a Twitter user name and monitor all incoming and outgoing public messages associated with the user. If several subjects of an investigation are identified as Twitter users, each of the profiles can be loaded in a separate column and monitored. Occasionally, this will result in two of the profiles communicating with each other while being monitored. You can also use the Geo search mentioned earlier within Tweet Deck. A column that searches "geocode:43.430242,-89.736459,1km" will display a live feed of Tweets posted within the specified range. A more precise search of "geocode:43.430242,-89.736459,1km fight" would add the keyword to filter the results. Figure 5.12 displays Tweet Deck with several searches.

The columns of Tweet Deck are consistently sized. If more columns are created than can fit in the display, the "Columns" option with left and right arrows will provide navigation. This allows for numerous search columns regardless of screen resolution. This is an advantage of Tweet Deck over the other services discussed. Tweet Deck is one of my Twitter staples. I use it at some point during every investigation. I recommend familiarizing yourself with all of the features before needing to rely on it during your searches.



Figure 5.12: A Tweet Deck search screen.

Hootsuite Feed (hootsuite.com)

While Tweet Deck is my preferred viewer of live Twitter information, it does not display well for large audiences. If I am broadcasting my screen to a digital projector for a room full of people to see, the text is usually too small to accurately view from a distance. Hootsuite offers a solution to this predicament. If you wanted to display a live feed of anyone mentioning "OSINT" on Twitter, you can navigate to the following website in your web browser, after connecting to a Twitter account.

https://hootsuite.com/feed/OSINT+Search

You can replace "OSINT" in the above address with any term or terms of interest. The result will be a live stream with a very large font that could be viewed from a long distance away. This view can be beneficial for situations where Twitter streams are monitored by a group of people in an operations center.

Twiangulate (twiangulate.com)

This is one of two websites that I recommend to quickly identify any mutual friends or mutual followers listed on two Twitter profiles. It identifies mutual friends on two specific accounts. In one example, 521 people were friends with one of my subjects. However, only 15 were friends with both targets of my investigation. This can quickly identify key users associated within an inner circle of subjects. All 15 subjects were listed within the results including full name, photo, bio, and location. While Twiangulate has assisted me in the past, I now recommend Followerwonk as a superior solution.

Followerwonk (followerwonk.com)

The second website that I use for group Twitter analysis is Followerwonk. This service offers more options than Twiangulate and will let you compare up to three users. The second tab at the top of the page, titled "Compare Users", will allow you a more thorough search. Figure 5.13 displays the analysis of three subjects. You can see that the first and second subject do not have any people in common that they follow on Twitter. This can indicate that they may not know each other in real life, or that they simply have different tastes in the people that they find interesting. However, the first and third subjects have 79 people in common that they follow on Twitter. This is a strong indication that they know each other in real life and have friends in common. Clicking on the link next to this result will display the identities of these people as seen in Figure 5.14.

This default search on Followerwonk is a good start. A more valuable search is to analyze the people that follow these users. The previous example identified people that our targets followed. This will often include celebrities, businesses, and profiles that probably have no impact on your investigation. However, the people that follow your targets are more likely to be real people that

may have involvement in your investigation. Figure 5.15 displays the results of the same targets when the search criteria was changed to "Compare their followers" in the dropdown menu next to the search button. We now see that the first and second subject still have no one in common. The first and third subject have 200 people that follow both of their Twitter feeds. You can click the result link to identify these 200 people.

Followerwonk possesses other search capabilities for user analysis. The first tab at the top of the screen will search any term or terms to identify any Twitter bios that contain those words. This may identify profiles that were missed during the search on twitter.com for messages. The third tab, titled "Analyze Followers", allows you to enter a single Twitter handle and either analyze the people the user follows or the people that follow that user. The second option usually provides more relevant results.

The information provided during this search will display numerous pie charts and graphs about the user. The most useful is the map that identifies the approximate location of the people connected to the person's Twitter account. Figure 5.16 displays a map for one of the users searched previously. This provides a quick indication of the regions of interest to the target. Figure 5.17 displays a detail level of an area near Denver. Each small dot identifies an individual Twitter account of a person that follows the target and lives or works in the general area. This location data is very vague and does not usually correctly correlate with the address on the map. This should only be used to identify the general area, such as the town or city, of the people that are friends with the target on Twitter. In the past, I have used this data to focus only on people in the same area as my homicide victim. I temporarily eliminated people that lived in other states and countries. This helped me prioritize on subjects that could be contacted quickly and interviewed.

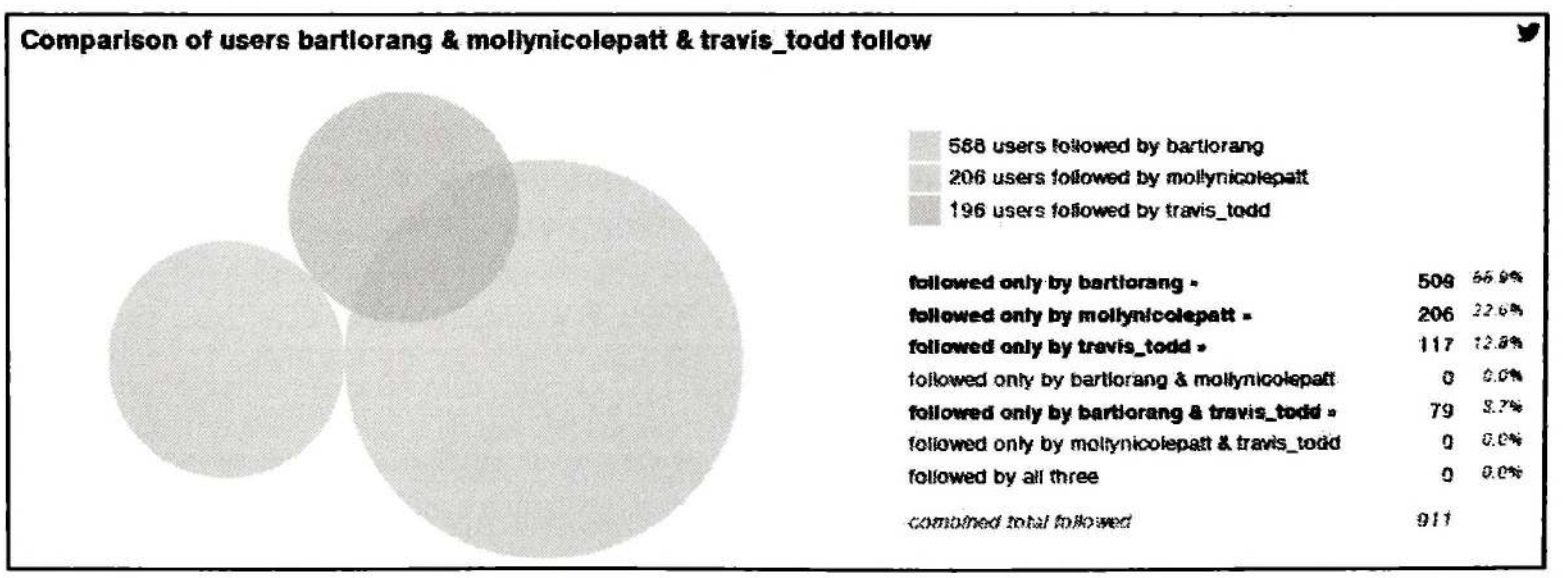


Figure 5.13: A Followerwonk user comparison.

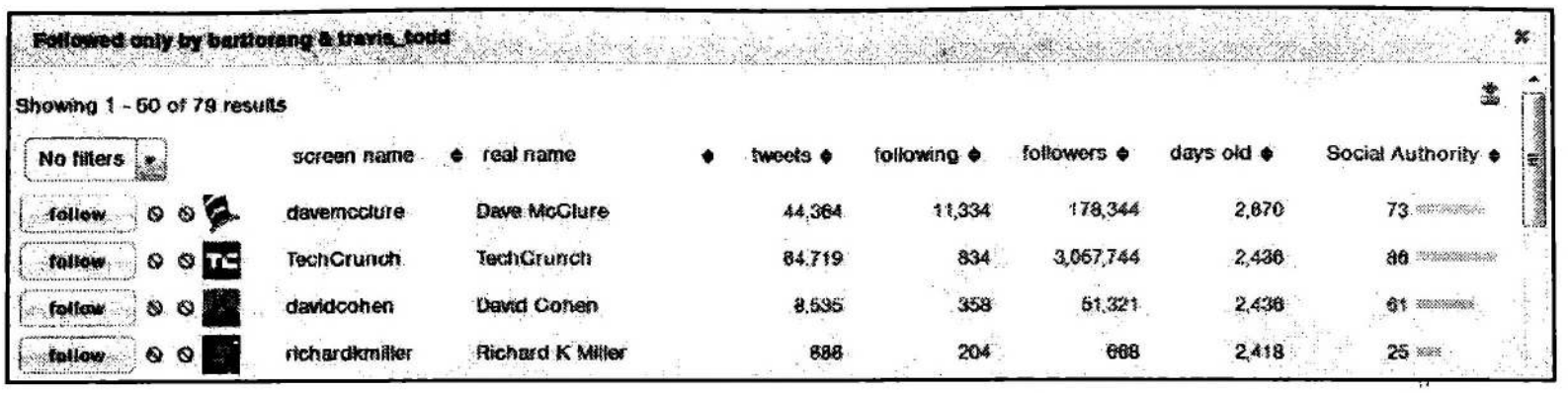


Figure 5.14: A Followerwonk list of users.

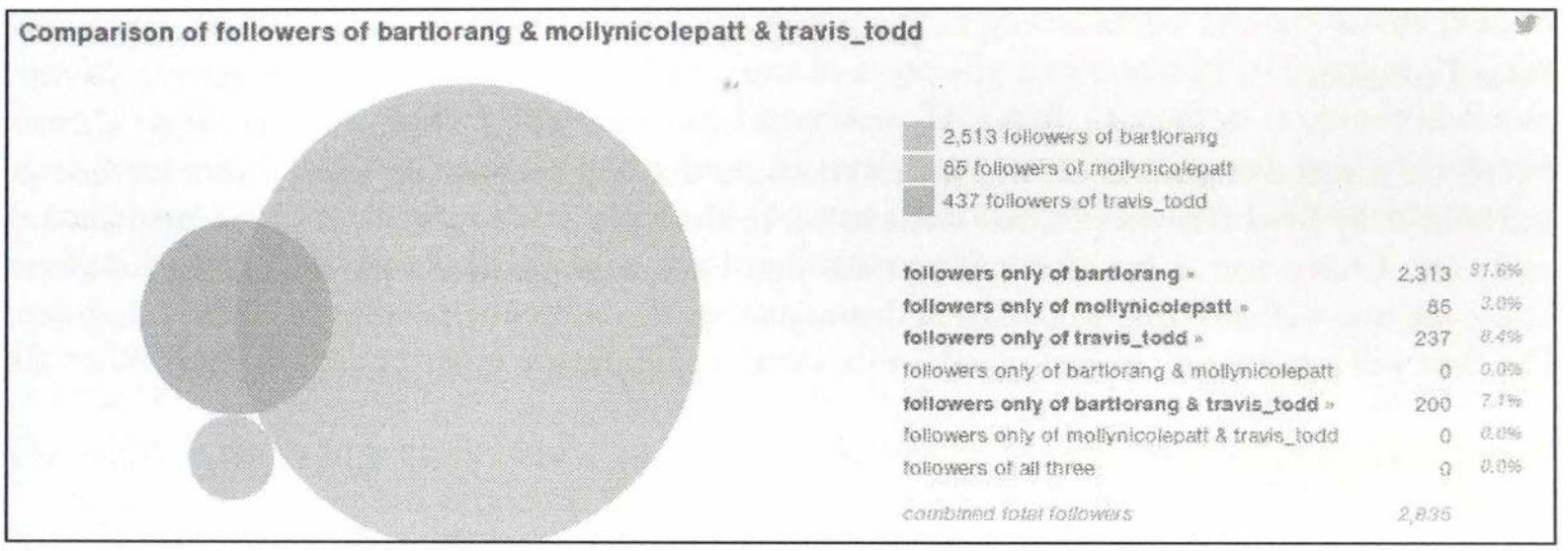


Figure 5.15: A Followerwonk user comparison.

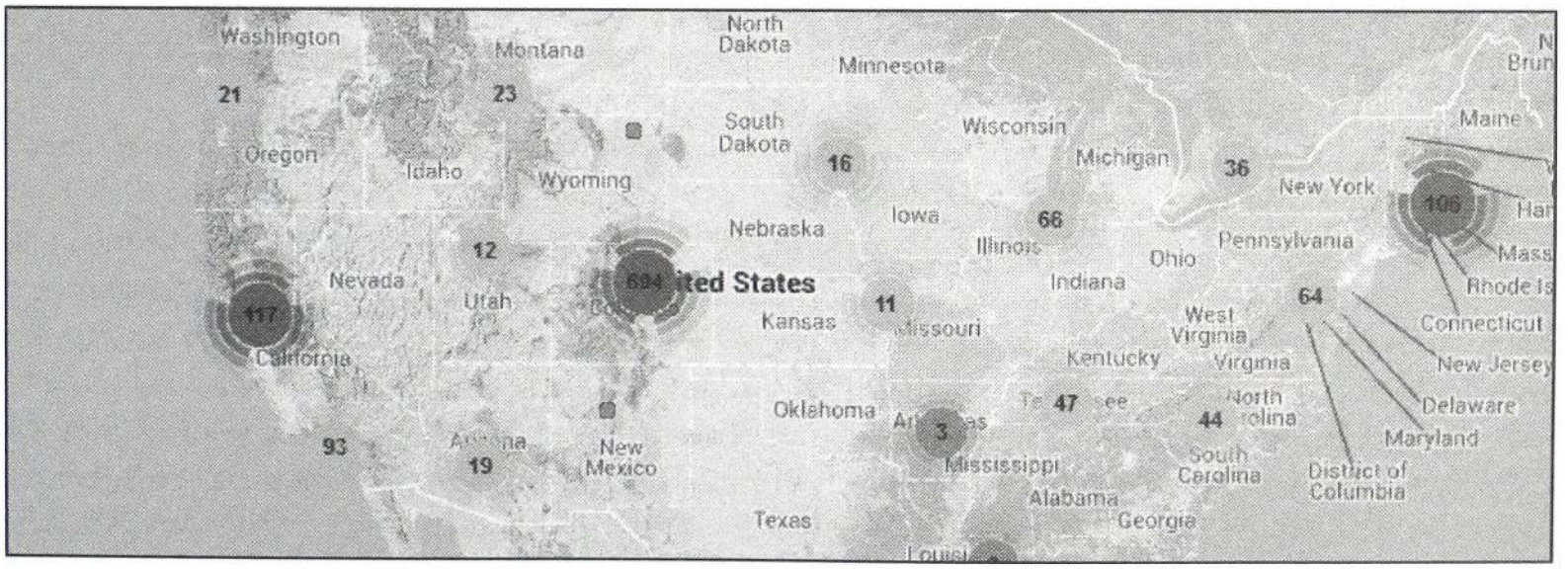


Figure 5.16: A Followerwonk map of users connected to a target.

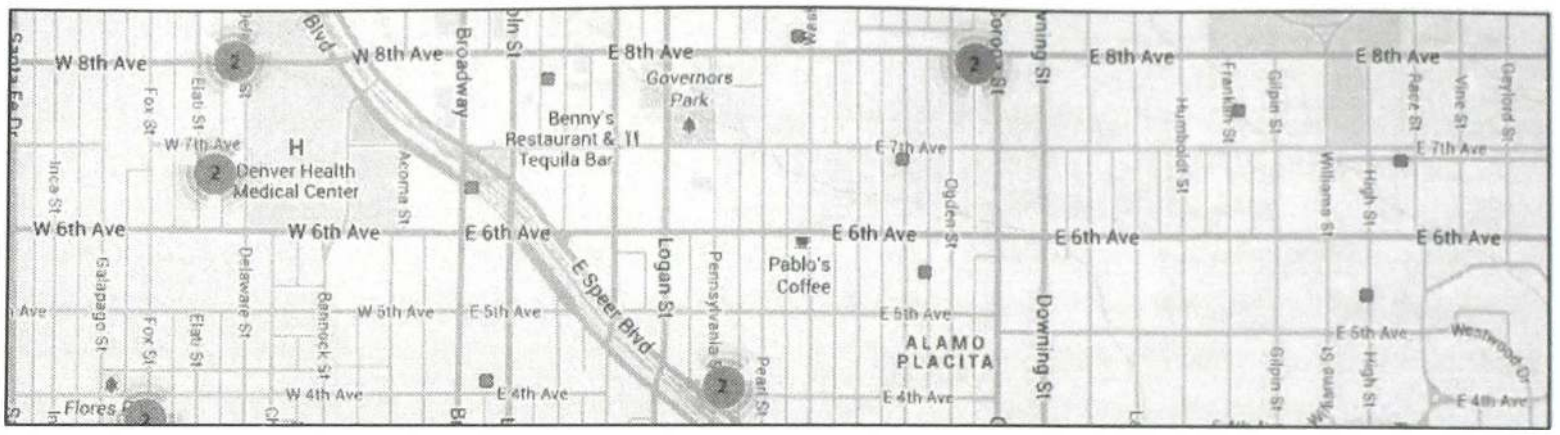


Figure 5.17: A detailed view of a Followerwonk map of connected users.

Fake Followers

There are a surprising number of Twitter accounts that are completely fake. These are bought and sold daily by shady people that want to make their profiles appear more popular than they really are. I have seen a lot of target profiles that have been padded with these fake followers. There are two websites that will assist in distinguishing the authentic profiles from the fraudulent. The first will require you to be logged into a Twitter account.

Status People (fakers.statuspeople.com)

The strength of this service is that it gives you additional information such as the languages spoken by the followers, the people that have not posted a Tweet in the past 100 days, and the users that follow less than 250 people. The weakness is that this is a paid service and it limits the free visibility. You are allowed to search only two Twitter handles during your free trial and you can see only two of the fake profiles identified. The website is poorly designed, and I would be skeptical of providing a credit card to the service.

Twitter Audit (twitteraudit.com)

This service does not require you to have a Twitter account, but the results are not as accurate. Each audit takes a random sample of 5000 Twitter followers for a user and calculates a score for each follower. This score is based on number of Tweets, date of the last Tweet, and ratio of followers to friends. They use these scores to determine whether any given user is real or fake. Of course, this scoring method is not perfect, but it is a good way to tell if someone with lots of followers is likely to have increased their follower count by inorganic, fraudulent, or dishonest means.

Miscellaneous Twitter Sites

Every week, a new site arrives that takes advantage of the public data that Twitter shares with the world. These sites offer unique ways of searching for information that Twitter does not allow on

their main page. This partial list is a good start to finding information relevant to your target.

BackTweets (backtweets.com)

BackTweets provides one unique service. It identifies any Twitter posts that included a link to a specified website. This tool works even if the post used a URL shortening service such as bit.ly. This can be used to locate subjects that are promoting illegal websites, followers of protest movements, or to identify the popularity of a website.

Trendsmap (trendsmap.com)

Monitoring trends on Twitter can provide intelligence on a global scale. The keywords that are currently being posted more than any other terms in a specific area could be of interest. This can identify issues about to surface that may need attention. This type of analysis is common during large events such as protests and celebrations. Several websites offer this service, but I choose Trendsmap. You can search either topics or a location. Searching a location will provide the top keywords being posted as well as a heat map to identify peak usage.

MentionMapp (mentionmapp.com)

This service displays a visualization of a user's Twitter connections and topics. It identifies connections to other users based on Twitter traffic. The width of the connecting lines identifies the strength of connection. This can provide a quick look at strong associations without analyzing an official Twitter feed. Note that this only extracts the most recent 200 Tweets and analyzes the communications within them.

Twitonomy (twitonomy.com)

One Twitter analytics website that stands out from the rest is Twitonomy. This is the most complete analytics service that I have found for a single Twitter handle. A typical user search would fill four pages of screenshots. A search of the user "humanhacker" immediately revealed the following details.

He has posted 8,689 Tweets
He is following 170 people
He has 17,079 followers
He joined Twitter on June 14, 2009
He averages 5 Tweets per day
He has mentioned 4,175 other Twitter users
He replies to 34% of posts to him
He has reTweeted 621 posts (15%)

The remaining sections of this page identify current posts, followers, people following, favorites, and lists. The main analytics portion identifies the average number of posts by day of the week and by hour of the day. It also displays from which platforms the user Tweets. Figure 5.18 displays this portion. This data discloses that the target has an Android and an iPhone, and that the majority of his Twitter time is spent on a Mac computer. This also identifies his preferred web browser, check-in utility, photo sharing service, and video sharing service. Other information includes his Tweets that are most "favorited" and "reTweeted", the users to whom he replies most often, and the users whom he mentions more than others. If you have a Twitter name of interest, I highly recommend searching it through Twitonomy.

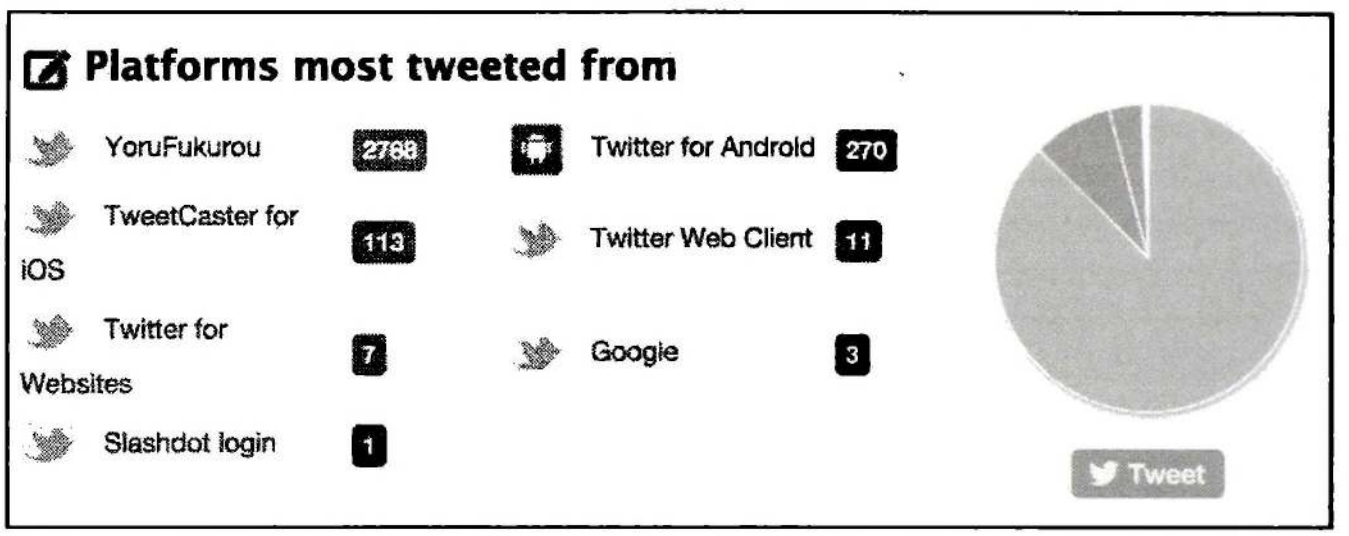


Figure 5.18: A portion of a Twitonomy search result identifying user platforms.

Tinfoleak (tinfoleak.com)

This Twitter analytics tool provides a simple yet thorough report. It was mentioned as an application in Chapter Two, but you can take advantage of a web option here. The website requires that you log into your Twitter account through this service, and the report includes the following information relevant to the target.

Twitter ID #	Location	Hashtags
Account Creation Date	Time Zone	User Mentions
# of Followers	Number of Tweets	Metadata from Images
# Following	Twitter Clients Used	Geo-Location Data

Note that using the Linux program inside Buscador does not provide your target data to the Tinfoleak website, while this web-based service does store target data on their web server. If you have a sensitive investigation, you should consider the Buscador option. As proof of this, consider the following URL which was created after I conducted a query of my own account.

https://tinfoleak.com/reports2/inteltechniques.html

FollerMe (foller.me)

This service is very similar to the previous Twitter analytics options. Providing a Twitter user name presents the typical bio, statistics, topics, hashtags, and mentions analysis that you can find other places. I find the following option of most interest to my investigations. I previously explained Sleeping Time as a resource to learn a target's usual sleep pattern based on posting times. FollerMe provides a bit more detail including posting patterns per hour. Note that the results are displayed in Universal Time, so you will need to convert as appropriate for your suspect. Figure 5.19 displays the results for my account. Since I am on the east coast (UTC -5), this example indicates that I tend to never post before 8:00 a.m. or after 11:00 p.m. My peak Tweeting time is at 11:00 a.m. There is a very obvious sleep pattern present within this result.

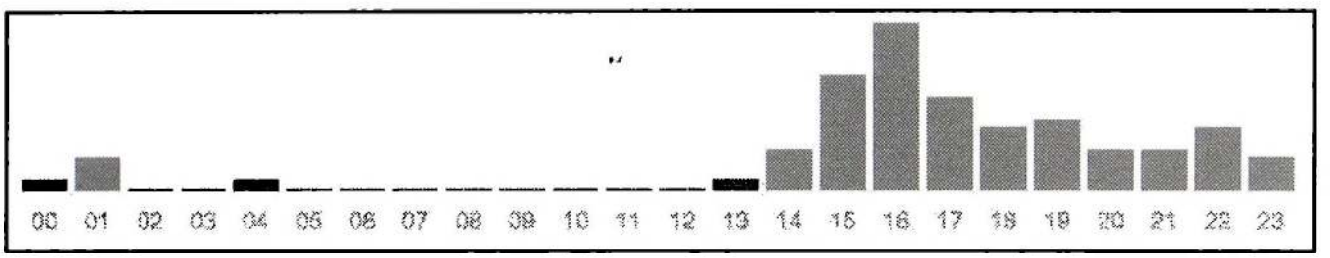


Figure 5.19: A FollerMe posting history by time.

TweetTopic (tweettopicexplorer.neoformix.com)

This simple tool provides one feature that I have found helpful in my investigations. Once you supply the target Twitter user name, it collects the most recent 3,200 Tweets and creates a word cloud. This identifies the most common words used within posts by the target. There are several sites that do this, but this service takes it a vital step further. Clicking on any word within the result displays only the Tweets that include the selected term. In Figure 5.20, you can see that I tend to Tweet about OSINT more than anything else. Clicking on any of the "ios" or "Facebook" circles would immediately identify posts related to those terms. I have used this when I have a target with too many posts to read quickly. TweetTopic allows me to quickly learn what my target posts about and immediately delve into any topics of interest to my investigation.

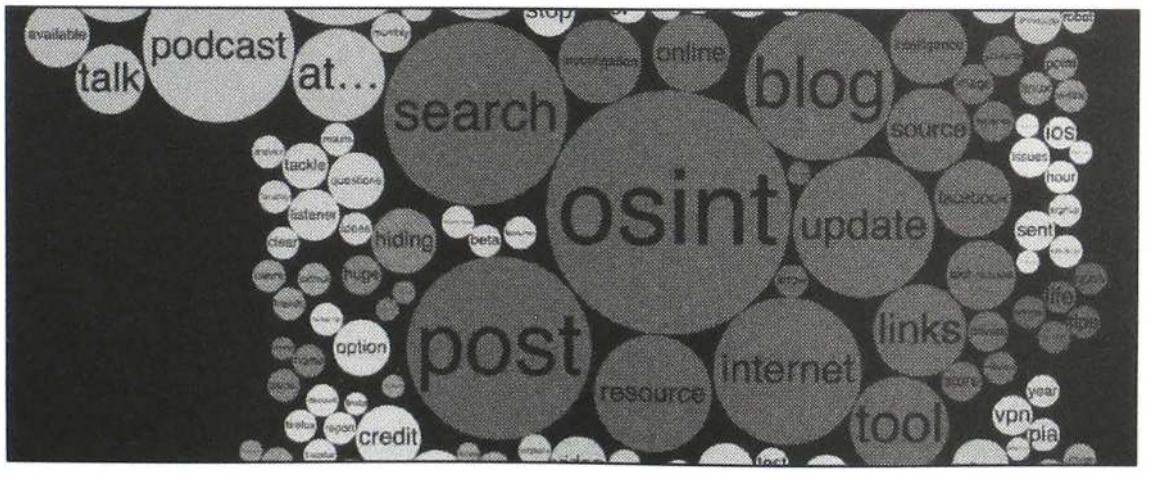


Figure 5.20: A TweetTopic interactive word cloud.

Real World Application: Parents can benefit from these Twitter searches. Locating a child's profile can provide great detail about portions of the child's life that are not shared with the parent. On one occasion, I was presenting to a group of concerned parents about the dangers of the internet. Afterward, a parent located her child's profile and discovered an abundance of Tweets that referenced depression and suicidal thoughts. The parent had no idea that her child was having such troubles. Professional assistance with the situation was sought. The parent could also identify a specific location that the child was visiting every day after school that was previously unknown to the parents.

On a final note about Twitter, I believe it is under-utilized by most investigators. We tend to find our target's profile, scroll a bit, and quickly dismiss it as a valuable resource. In my experience, the bulk of relative information, and the most valuable to my case, is never present in this live view. We must uncover the less obvious evidence in order to thoroughly investigate our subjects. The next time you face a Twitter profile, I hope these resources take you in a direction that most would not consider.

CHAPTER SIX
SOCIAL NETWORKS: OTHERS

Facebook and Twitter are likely going to offer the most bang for your buck in terms of web-based social networks. However, there are many other options, and all should be researched. While we see high usage on the most popular sites, the vast majority of the content will be of little interest to you. Instead of finding your suspect confessing to your investigation on Facebook or Twitter, you may be more likely to find his grandmother testifying to his innocence. Smaller networks often provide more intimate details. Your suspect may feel exposed on the bigger networks, but a bit more private on smaller websites. This chapter begins with the largest of the remaining networks that will be discussed, and ends with a collection of lesser-known options. Note that mobile-only networks will be discussed later in Chapter Twenty-One.

Instagram (instagram.com)

Instagram is a photo-sharing service that is now owned by Facebook. With well over 800 million active monthly users, the amount of content available here is overwhelming. This application works alone or in correlation with Facebook and Twitter to distribute the photos. This service is very popular with most photo sharing Twitter users and should not be ignored. Searching for Instagram content through Twitter's website will not provide all of the content. Surprisingly, there is no search feature on the Instagram home page. If you want to use Instagram's search database, you must connect directly to an account such as Instagram.com/mikeb. However, this search field only identifies users and hashtags related to the search terms. It does not provide a true keyword search. We will use Google for this in a moment.

In previous editions of this book, I detailed several third-party Instagram search options that unlocked a lot of hidden content within user accounts. On June 1, 2016, Instagram tightened their API, and this killed most of the useful websites. My own Instagram tools page suffered drastically from the new restrictions, and I had to start over with new options. Fortunately, you still have many search options for your next Instagram investigation. Let's start with keyword searching.

I have found greater success with a custom Google search instead of an Instagram search field. The following query on Google will produce 53 results that display Instagram posts that mention "OSINT" within the post title or comments.

site:instagram.com "OSINT"

This same term searched on an Instagram page only displayed users that have the term within the user name. When searching "#OSINT" on Instagram, I was provided a list of hashtags that include the keyword. Each of these hashtags are associated with multiple Instagram posts.

Consider the following example in order to identify the benefits of searching away from the Instagram website. While on an Instagram page, I searched for my target Alex Albrecht, which presented several user names which may be related. One of these was my target, "alexalbrecht", and we can analyze that profile page later. Similar to Facebook and Twitter, a person's Instagram profile only tells a small portion of the story. We have reached the limit of usefulness on the official Instagram page. Conducting the following search on Google revealed hundreds of results.

site:instagram.com "alexalbrecht"

These are the various pages and posts that contain the text of my target's user name. Many of these are posts from the target, which we have already seen by looking at his profile. I prefer not to see these redundant results, so I conducted the following search on Google.

site:instagram.com "alexalbrecht" -site:instagram.com/alexalbrecht

This tells Google that I only want to search Instagram, I want the user name of alexalbrecht to be present on all results provided, but I do not want to see any results that begin with instagram.com/alexalbrecht. This prevents results from our target's own profile. The remaining results include other people's pages that mention our target, and comments posted by our target. The next search can filter a bit more by only locating Instagram posts that "tag" our target within the message.

site:instagram.com "@alexalbrecht" -site:instagram.com/alexalbrecht

Similar to the previous search, it forces Google to ignore the target's profile, and only focus on people that are posting "to" the target with @alexalbrecht instead of just his user name alone. These searches can be modified to include or exclude keywords, user names, real names, or any other terms with which you have interest. You could also repeat these on Bing and Yandex for better coverage. Hopefully, you now have a target profile of interest.

Instagram Profile Image

There is one obstacle with Instagram searches that needs to be addressed. When you identify a profile of interest, there will always be a profile image at the top of the page. This is usually a 150x150 pixel image that is small and heavily compressed. In 2015, Instagram changed the way that these thumbnails and original images are stored. The goal of this technique is to locate these small thumbnail images and identify the full high-resolution images that were originally uploaded. For an example, I will pick on the user "Johnny". His profile is at instagram.com/johnny. When I right-click on his profile picture, and choose to open the image in a new tab, I am presented with the following URL that displays a very small photo of the user.

https://scontent-sjc2-1.cdninstagram.com/hphotos-xtp1/t51.2885-19/s150x150/917360_1513292768967049_387615642_a.jpg

Note the portion that displays s150x150. This is telling Instagram to retrieve the 150-pixel thumbnail for the user and display it. If you remove that portion of the URL, you will receive a much different result. Instead, navigate to the following URL.

https://scontent-sjc2-1.cdninstagram.com/hphotos-xtp1/t51.2885-19/917360_1513292768967049_387615642_a.jpg

The result is the original image at full resolution. The detail of this image trumps the original thumbnail presented with the profile. Note that this technique only works on users that have updated their profile image since early 2015.

Instagram Private Accounts

There is no known way of displaying Instagram posts within a user's profile if it is set to private. However, one technique might be beneficial if a user posts through other social networks. As an example, consider the Instagram user shoegirlcorner at instagram.com/shoegirlcorner. This private account only reveals a profile photo and general details about the user. The posts are all blocked. Using the previous technique, we can view the entire profile image. If you conduct a reverse image search of this photo (explained in Chapter Fourteen), it connects to a LinkedIn profile of Loida Casares. A Google search of Twitter Loida Casares identifies this user's Twitter account. Visiting that page and searching for Instagram reveals several posts where the user posted a photo on Twitter that connects to an Instagram account. Clicking any of these images reveals the entire post. While you can never connect straight to these images through a private Instagram profile, many users post these images through public Twitter feeds. The Instagram post is not private, only the profile page possesses the restriction. Another way to accomplish this is through the following search on Google, Bing, or Yandex.

site:twitter.com "user name" "instagram.com/p"

This tells Google to only search on Twitter, mandates that the supplied user name is present, and forces "instagram.com/p" to be included within the content. The last rule is because any Instagram images posted to Twitter will always have a hyperlink title that starts with the domain followed by "/p". Note that the user name could be either a Twitter handle or an Instagram user.

Once you have located your Instagram content of interest, you need to archive your evidence. Chapter Two presented an automated solution for bulk downloading media, and Chapter One identified ways to capture the profile views through browser extensions. If a photo is linked through Instagram in a Twitter post, you should download the digital version with the best quality available. Any links should open a new tab that will load the photo within the Instagram page. This photo will not have any options to download or save the data. If you right-click on it, you will not be given an option to save-as or download the image as you would on standard websites. Instead, we must manually archive this file. The process for archiving an embedded photo in Instagram will vary based on your browser. The easiest is with Firefox. Right-click on the image

of interest and select "View background image" or "View Image Info". This will present the actual jpeg image, which can be saved by right-clicking again and choosing "Save image as".

You should also consider digging into the source code of your evidence in order to identify further details that could be valuable. First, I like to identify the user number of an account. Similar to Facebook and Twitter, people can change their user name on Instagram, but not their user number. Right-click on your target profile page and select the option to view the source code. This will open a new tab with a lot of pure text. With either ctrl-f or command-f, conduct a search for exactly the following. The numbers directly after this data will be the user number of your target. In our previous example, his user number is 307949.

"owner": {"id":

While we are looking at the source code of pages, we can use this technique to identify the exact time that a post was published. Instagram only identifies the date of a post, and not the time. This detail can be crucial to a legal investigation. For this example, I located an image posted by our target at the following address.

https://www.instagram.com/p/Bcivg7BjJb3

Viewing the source code of this page, look toward the 70th line of code for a section that begins with "meta property="og:title"". The following is the entire line, and we now know the exact time, in universal time format.

<meta property="og:title" content="Instagram post by Alex Albrecht Dec 11, 2017 at 1:04am UTC

Followers & Following

Instagram now requires users to be logged into an account in order to view the followers of a target or the profiles that a target follows (friends). Fortunately, Instagram is fairly forgiving with new profile creation, and the instructions at the beginning of Chapter Four should help you create your own investigation account. Viewing these lists is not a challenge, but proper documentation can be tricky. The following explains how I choose to view and document every Instagram follower and friend of my target account, using timferriss as an example.

After logging into your account and navigating to the target profile, you will be able to simply click on "Followers" or "Following". In this demonstration, I chose the people he is "following", often referred to as friends. This opened a new window on my screen with the first twenty people displayed. Since timferriss follows 196 people, I can either scroll down the entire list or press and hold the space bar to load all of the accounts. You will not be able to see all of the accounts at once because the window only displays ten within its boundaries. This causes screen capture tools to be useless in this scenario. If you are using either the Firefox browser within Buscador (Chapter Two) or your own version of Firefox with the recommended add-ons (Chapter One), right-click

within this window and choose "Copy All Links" > "Current Tab". This copies all of the Instagram account hyperlinks within the target's followers list into your computer clipboard. Open your desired documentation software, such as Word, Excel, or any text editor, and paste the contents into the application.

I prefer Excel (or Calc within the free LibreOffice suite), because I want to easily delete duplicates. Each user is likely mentioned twice within your results. This is due to a hyperlink from their name, and an additional link from their profile photo. While in Excel, select the column where you pasted the results. Click on the "Data" tab and choose "Remove Duplicates". Column A will now only contain one URL per account. In Calc, select the column where you pasted the results (A). Click on "Data" > "More Filters" > "Standard Filter", then expand "Options". In the top right "Value" field, choose "Not Empty". Select "No duplications" and "Copy results to". Enter "B1" below "Copy results to" and click "OK". Column B should now only contain one result for each account. Below is an excerpt of this data.

185	https://www.instagram.com/modernoutdoors/
186	https://www.instagram.com/jimmahfood/
187	https://www.instagram.com/jorge_jimenez_art/
188	https://www.instagram.com/kromninja/
189	https://www.instagram.com/wilderness_culture/

In my experience, Instagram limits the number of accounts within the window to 1,000. This should be sufficient for most friends of the target, but may be limited when looking at the followers of famous people. If you want to obtain a single screen capture of all of these accounts, I recommend **SomeTag** (sometag.com). Searching your target account within this service and selecting either Followers or Friends loads the same data within one screen. This is fine for visual capture, but not for extracting hyperlinks. While each link result will display the Instagram user name and number, the URLs connect directly to SomeTag pages. The first friend of timferriss is linked to sometag.com/account/iamjamiefoxx/256432578/. I believe these two methods can complement each other, and neither provides a perfect solution.

Third-Party Tools

Overall, the various Instagram search websites do nothing more than what we replicated in the previous instruction. However, these sites could identify that one piece of vital evidence that we missed. Therefore, it is always best to have as many resources as possible. I have had limited success with **Websta** (websta.me) and **Lakako** (lakako.com).

IntelTechniques Instagram Search Tool (inteltechniques.com/osint/instagram.html)

If this seems like a lot of work for minimal information, consider using my custom Instagram search tool. Figure 6.01 displays the current version of this page, which includes the search options previously discussed. In this image, you can see that I have conducted the previous two source-code search methods, and the results appear within the page. This simplifies the process.

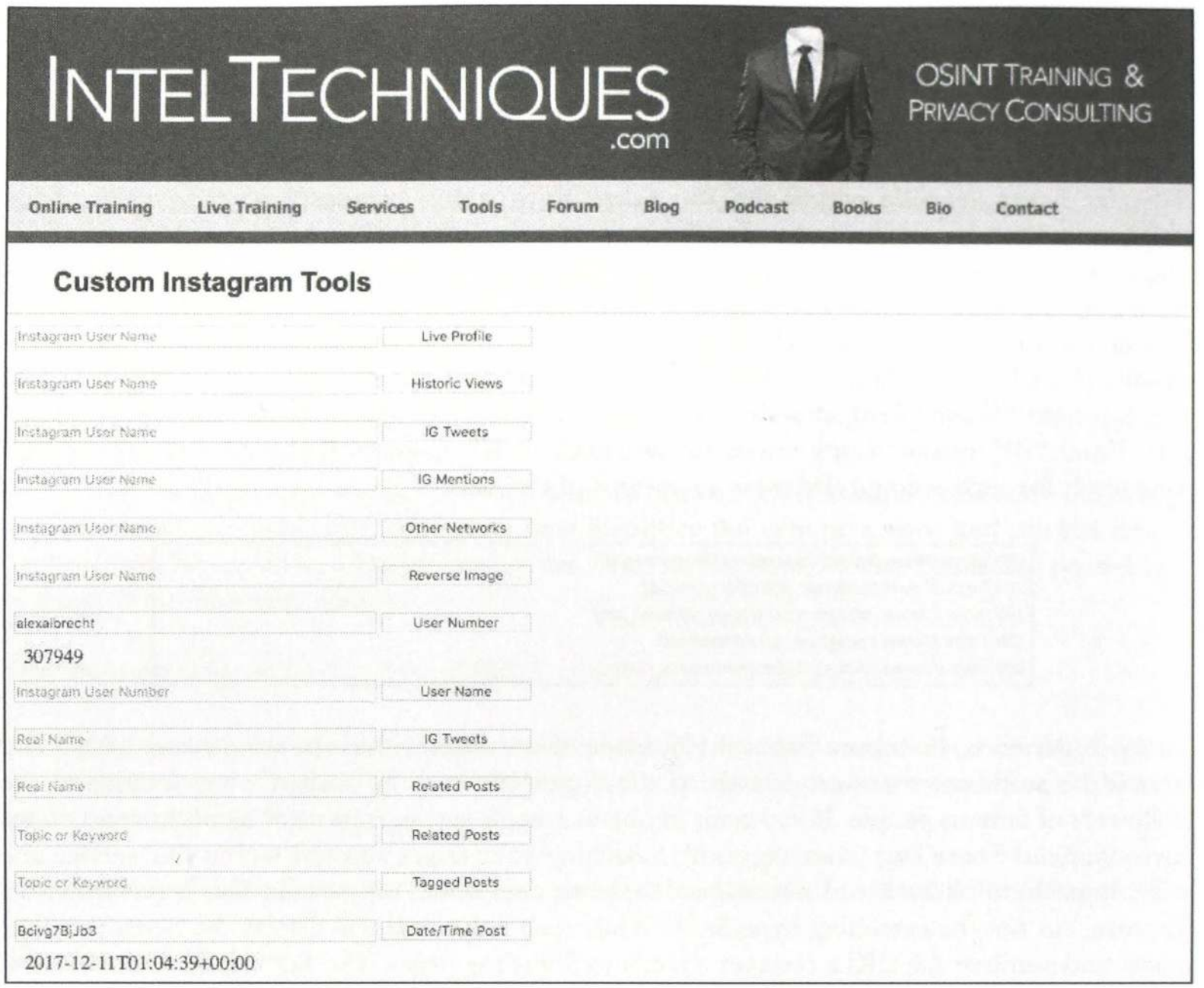


Figure 6.01: The IntelTechniques Custom Instagram Search Tool.

Google+ (plus.google.com)

Google's social network is fairly straightforward. All Google+ pages have a search bar at the top ready for a real name to search. While Google has allowed the use of screen names in rare circumstances, a real name is required by the service. After locating the profile of your target, you can click through the options for Posts, About, Photos and Videos. These pages will display information about the target that was supplied by the user. The left column will display a photo posted by the user and an abbreviated list of people involved in "circles" with the user. These circles are Google's way of identifying relationships between users. The right column will identify additional profiles on other networks of the user as well as links uploaded by the user. I know of no techniques that will obtain further information than what is readily available through simple searching.

LinkedIn (linkedin.com)

When it comes to business-related social networking sites, LinkedIn is the most popular. It is owned by Microsoft and currently has more than 467 million subscribers internationally. The site requires searchers to create a free profile before accessing any data. As with any social network, I recommend creating a basic account with minimal details. The search field of any page offers a search using a real name, company, location, or title. These searches will often lead to multiple results that identify several subjects. The upper right portion of this results page will offer some basic refinements to the search to filter by name, title, company, school, location, and others. Knowing the real name will be most beneficial. The results page will include the target's employer, location, industry, and a photo. After identifying the appropriate target, clicking the name will open that user's profile. If searching a common name, the filters will help limit options.

The profiles on LinkedIn often contain an abundance of information. Since this network is used primarily for business networking, an accelerated level of trust is usually present. Many of the people on this network use it to make business connections. Some of the profiles will contain full contact information including cellular telephone numbers. This site should be one of the first stops when conducting a background check on a target. The target profile often contains previous employment information, alumni information, and work associates. Aside from searching names and businesses, you can search any keywords that may appear within someone's profile. Since many people include their phone numbers or email addresses in their profile, this can be an easy way to identify the user of that specific data. Visiting this profile identifies further information as well as confirmation of the target number.

Searching by Company

If you are searching for employees of a specific company, searching the company name often provides numerous profiles. Unfortunately, clicking on any of these profiles presents a very limited view with the name and details redacted. The name of the employee is not available, but the photo and job description are usually visible. You are now required to upgrade to a full premium account, or be in the same circles as the target, in order to get further information. Instead, consider the following technique.

Search for the business name of your target company, or the employer of your target individual. I typed "Uber" into the search bar and received the official business page on LinkedIn. Clicking the "See all 41,220 employees on LinkedIn" link presented me with numerous employee profiles such as those visible in Figure 6.02. Notice that the names are redacted and only "LinkedIn Member" is available. Clicking this first result prompts me with "Profiles out of your network have limited visibility. To see more profiles, build your network with valuable connections". We struck out, but there are ways that you can proceed in order to unmask these details.

First, copy the entire job description under the "LinkedIn Member" title. In this example, it is "Account Executive at Uber". Use this in a custom Google search similar to the following.

site:linkedin.com "Account Executive at Uber"

The results listed will vary from personal profiles to useless directories. Since Uber is such a large company, I had to view many pages of results until I identified my target. When I opened the 24th search result, the LinkedIn page loaded, and her photo confirmed it was the correct target. The easier way would have been to search the images presented by Google. After the above search is conducted, click on the Images option within Google and view the results. Figure 6.03 (left) displays a section, which easily identifies the same image as the LinkedIn target. Clicking this will load the profile page with full name and details.

Another way to accomplish this is to navigate through the profiles in the "People also viewed" column. These pages include other profiles viewed that are associated with whichever person you are currently analyzing. These people may not be friends or co-workers with your target, but there is a connection through the visitors of their pages. As an example, I returned to the Google search at the top of this page. I clicked on the first search result, which was not my target. However, in the "People also viewed" area to the right, I saw my target, including her full name and a link to her complete profile. Figure 6.03 (right) displays this result.

Finally, the last option is to conduct a reverse image search on the photo associated with the target's profile. Full details of this type of search will be presented later. For this demonstration, I will right-click on her photo and choose Search Google for Image from my Chrome browser. While the first result is not the target, clicking the page does present a link to the target's unmasked page.

Searching by Country

While LinkedIn is an American company, it is a global social network. If you know that your target is in a specific country, you can filter your search accordingly. This can be done manually by navigating to a foreign subdirectory such as uk.linkedin.com (UK), ca.linkedin.com (Canada), or br.linkedin.com (Brazil). This tedious method of searching can be replaced with a custom Google search engine. Navigate to inteltechniques.com/osint/linkedin.country.html. You will see a single search field. Conduct your query and you will be presented with a new window that will filter your search results on LinkedIn by country.

PDF Profile View

You may want a quick way to collect the publicly available details from the profiles that you find. One option is to have LinkedIn generate a PDF of the information. While on any profile, click the three dots next to the name and choose "Save to PDF". This will not extract any private details, but will make data collection fast.

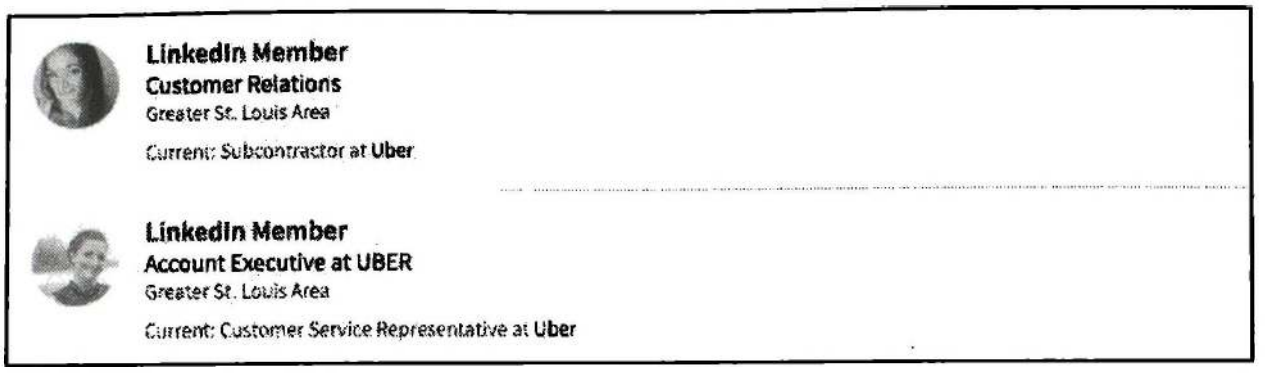


Figure 6.02: Redacted employee results from a business search.



Figure 6.03: Google Images results (left) and un-redacted LinkedIn results (right).

IntelTechniques LinkedIn Search Tool (https://inteltechniques.com/osint/linkedin.html)

As an attempt to simplify the advanced search options within LinkedIn and through third-parties, I created a custom online tool. Figure 6.04 displays the current version. The left column allows queries through a generated LinkedIn URL, isolating keywords, name, title, company, and education. This replicates the options available on the official site. The right column uses the same data to create a custom Google search in attempts to find direct profile URLs of your targets. The middle section converts a user name to the live profile or online images; searches a company name for employees and their photos; and conducts a reverse-image search on any profile pictures that you locate. The final tool allows you to combine multiple parameters and optionally filter by a keyword. As with other IntelTechniques tools, this one does not extract any information that you could not locate manually. Hopefully, it makes the process easier and faster.

Recruit'em (recruitin.net)

Recruit'em is a tool that helps you find people on social networks. It creates a basic Boolean string that searches for publicly available webpages using Google. It targets LinkedIn, Google+, GitHub, Xing, and others. You can filter by country, job title, location, education, and keywords. While using detailed Google operators could replicate the results, this simple interface may be preferred by most users. It was designed to be used by recruiters looking for new talent, but investigators can use it to locate profiles when a name search is inappropriate.

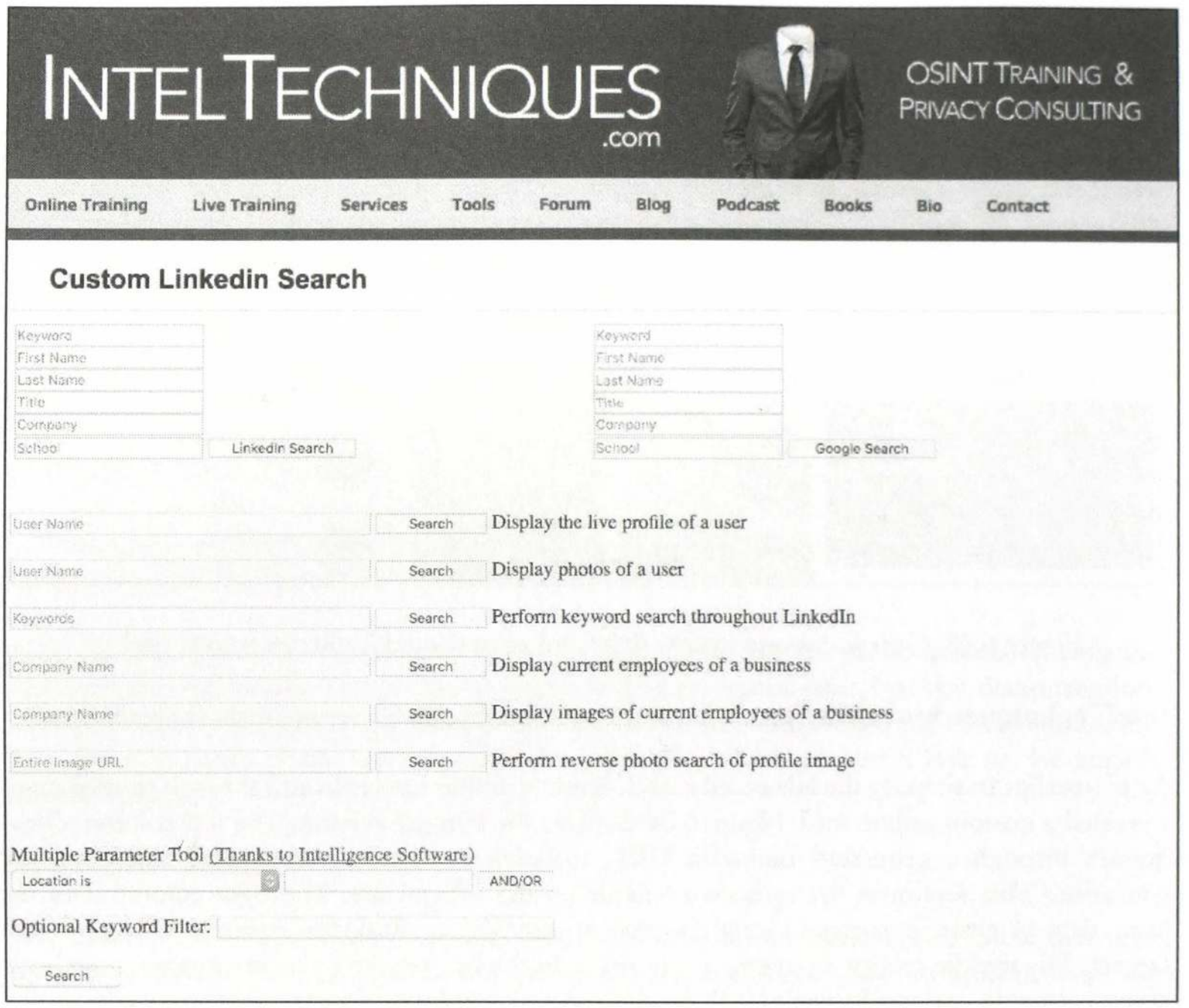


Figure 6.04: The IntelTechniques LinkedIn Custom Search Tool.

Tumblr (tumblr.com)

Tumblr was purchased by Yahoo in 2013. It is half social network and half blog service. It is gaining a lot of popularity with many active users. At the time of this writing, there were 373 million blogs and hundreds of billions of posts. These posts can include text, photos, videos, and links to other networks. The search method and layout is no longer user-friendly. The search feature will only identify blogs that were specifically tagged by the creator with your search terms. I suggest using the Google custom search engine described in Chapter Three. As an example, I conducted a search of "bazzell" within the official Tumblr search and received three results. I then conducted the following search in Google and received the appropriate 205 results.

site:tumblr.com "bazzell"

Snapchat

Overall, Snapchat is a difficult OSINT resource. It is only available officially as a mobile application, and there is not any sort of native web search. The majority of the content is set to auto-destruct after a specific amount of time, and is privately delivered from one user to another. While there are some extended search options within the app itself, I will only focus here on web resources. The most powerful engine I have located is **Snapdex** (snapdex.com). Enter a user name within the search option or navigate directly to snapdex.com/username and view the bio, a profile picture, location data, and public "snaps". This will only be useful if you know your target's Snapchat name, as any additional search features appear missing. Additional search options include **Snap Code** (snapcode.net) and **Snap VIP** (snapvip.io)

MySpace (myspace.com)

Previous editions of this book possessed a large area devoted to MySpace. It discussed ways to view hidden videos, photos, and comments from a target's profile. In 2013, MySpace redesigned the entire website and all of the original techniques were disabled. We could still search for information, but the methods were much more traditional. The majority of user profiles were marked as private and no content was viewable. In 2015, a new option surfaced that would again allow the display of private video files, but it was disabled later that year. MySpace has become a hub for musical artists and possesses very little content valuable to investigations.

Access to private profiles

There have been several "hacks" in the past that would allow a type of "back door entry" into a profile that is marked as private. By the time these methods become known, the vulnerability is usually corrected by the social network host. Websites or applications that publicly claim to be able to access this secured data are most often scams or attempts to steal your own passwords. In my experience, it is best to avoid these traps and focus on finding all available public information. At the time of this writing, a Windows application had recently surfaced that claimed to be able to obtain all information from within private Facebook accounts. The program did not work, but installed a malicious virus instead. If it seems too good to be true, it probably is.

Contact Exploitation

Earlier in Chapter Four, I explained how to add cell phone numbers as contacts to a Yahoo account in order to supply them to Facebook. When Facebook received the numbers directly from Yahoo, it believed that the contacts were "friends"; therefore, it identified the names and accounts associated with each number. I refer to this technique as contact exploitation, and the Facebook/Yahoo connection is not the only option for this type of activity. The following details identify the ways to use one service to get information from another. Basically, by adding details to one service, we can convince another to provide information that we would not otherwise have the authority to view. A few demonstrations should help explain.

Twitter does not support native search by email address or cellular telephone number. However, it will connect to an authenticated Gmail account and search for Twitter users within the contact list in the Gmail account. This is not as straightforward as the Facebook/Yahoo technique. In order for Gmail to release the contact information, email communication must have occurred between you and the target. In a controlled environment, I conducted the following test. I added the email address of Justin Seitz to a new covert Gmail account. I then requested my new Twitter account to search my Gmail contacts attempting to "Find Friends", but received no results. I then sent a blank message to Justin's email account and repeated the "Find Friends" attempt within Twitter. Justin's account was displayed. Apparently, if you send a message to any email address from a Gmail account, and then ask Twitter to check that account for friends, it passes these email addresses to Twitter for cross-reference.

Gmail is closely connected with Google+. If you hover over any email address within your sent or received messages, Gmail searches that contact within the Google+ database. You can then click on the profile icon present in the pop-up window to navigate directly to the related profile. While email address searching is supported on Google+, some queries fail for unknown reasons. If you have communication with the target within a covert Gmail account, even if it is a sent message without a response, it qualifies for connection to Google+ and considers the target within your "Circle".

This technique works throughout several social networking environments. I keep covert Yahoo and Gmail accounts solely for adding my target's contact information and asking networks to find friends based on this data. I am often presented with profiles in my target's true name as well as alias accounts. Many people are tricky enough to create an alias profile, but too lazy to connect it to a covert email address or Google Voice number. Chapter Twenty-One will explain ways to use this technique within an Android emulator for further exploitation.

We now know that locating someone's social network profile can reveal quite a lot about them. Just knowing a target name can make use of the people search engines that will identify places to seek more information. Unfortunately, sometimes the investigator does not know the target's name. A common scenario is an investigation about some type of event. This could be a specific violent crime, a bomb threat on a campus, or inappropriate chatter about a particular business. All of these scenarios require search engines that monitor social network traffic. There is an abundance of these types of services. Some will work better on reactive investigations after an incident while others show their strength during proactive investigations while monitoring conversations. Many of the sites mentioned here will find the same results as each other.

Overall, some of the strongest methods of searching social network traffic have already been discussed in the Facebook and Twitter chapters. Searching for traffic at the source, such as on twitter.com, will usually provide more accurate and updated content than an aggregated website of multiple sources. Furthermore, searching specific services through Google or Bing may sometimes quickly locate results that would be difficult to obtain anywhere else. The use of the site operator explained in Chapter Three will take you far. Aside from direct searches on social

networks and targeted search engine queries, there are other options. The accuracy of the services mentioned in the rest of this chapter varies monthly. Hopefully, you will find some of these websites to have value in your investigations. I believe that the options in this chapter should be used to supplement, not replace, the results obtained from previous methods.

Custom Search Engines

Chapter Three discussed the creation of custom search engines on Google. Two of the final products created were the Social Networks Search Engine and the Smaller Networks Search Engine. Both of these offer a quick and thorough search of both popular and lesser known social networks. They can often identify communication involving your target. I have found these to be the best initial searches for general content about a specific topic, user name, real name, or event. Both of these engines can be found on the IntelTechniques website under Links > Custom Search Tools. Additionally, they can be accessed at the below URLs.

Social Networks: https://inteltechniques.com/OSINT/social.networks.html
Smaller Networks: https://inteltechniques.com/OSINT/smaller.social.networks.html

Social Searcher (social-searcher.com)

I had previously discouraged users from attempting searches on the first version of Social Searcher. Since then, I have begun to rely on their free service to digest data located on the main social networks. You can provide any keywords, user names, or terms and receive the most recent results from Facebook, Twitter, Google+, and the overall web. It allows email alerts to be created for notification of new content matching your query. One of the unique features of this website is the free ability to export search results into CSV format. This output contains the Twitter user name, date & time, and entire message among other information. Having this in a spreadsheet format can be incredibly beneficial. This document also included dozens of Reddit and other network posts. The document could easily be imported into any other data collection system.

Social Mention (socialmention.com)

Social Mention searches much of the same traffic as most of the other social search sites. This is one of the few sites that offer real-time statistics within the search results that can be beneficial to a researcher. These new sources of information include a sentiment reading, a passion reading, an average time frame per comment, and the top keywords present for the search conducted. This data will notify the researcher if the overall results for the search are negative, positive, or neutral. Furthermore, the identification of the most used keywords may provide further intelligence about additional terms that should be searched. Overall, I have found Social Mention to extend the sources of data to include personal blogs, photo sharing sites, and news.

Account Export Options

The following utilities can be very useful during investigations. Consider a scenario where you have cooperation from a target and consent to view social network accounts. Many suspects will allow you to peek into their online activity with hopes that it will stop your suspicion about them as a suspect. If you have explicit consent, consider collecting all available content from within the target's profiles. Up to this point, I have explained how to properly collect data from the open and public internet. This does not include a person's email content, calendar entries, truly private photos, or personal communication. I believe that any time you have permission from the target to view their accounts with their volunteered credentials, you should also request to archive the contents. This can be very difficult if done manually. The following techniques identify the easiest and most automated solution for the most popular environments.

Google Takeout (takeout.google.com/settings/takeout)

If your target has a Google account, there is an abundance of data available in several different areas. This can include Gmail messages, YouTube channels, Blogs, Calendars, Contacts, and many others. While logged in as the target, navigate to the above website. By default, every option should be selected. Clicking "Next" will forward you to a download page. Accept the default settings and click "Create Archive". Google will package every possible piece of data from the user's account and present it in very large compressed zip files. These can be opened or stored for later view.

Facebook (facebook.com/help/131112897028467)

Facebook does not offer a specific page for archiving data, but the feature is embedded into the user settings for every account. While logged in as the target, click the top right menu of any Facebook page and select "Settings". Click "Download a copy of your Facebook data" below the General Account Settings tab. Click "Start My Archive". Facebook will send an email to the address on file for the target. A download link in that message will present a compressed file of the entire contents of the target's Facebook profile. It is important to have consent on the target's email account in addition to Facebook for this method to work.

Twitter (support.twitter.com/articles/20170160)

Similar to Facebook, Twitter allows you to export an entire account when authenticated as the target user. Click on the profile icon in the upper right area and choose "Settings". On the left menu, choose "Your Twitter Data" at the bottom. Scroll to the bottom of the page and click "Twitter Archive". Click the button labeled "Request Your Archive" and the process will begin. A link will be sent to the email address on file for the target, and it will connect to a compressed file containing the entire account.

During my criminal investigations, I always asked for written consent to view and collect data

within my suspect's accounts. When allowed, I would ask them for his or her password and advise that I would go start the process of viewing and collecting the data. I would then execute the archive collection process and return to the interview. This way, downloading the data was consensual, and the target can withdraw consent at any time. Practically every popular service allows users to export their own data. Searching the provider and "export my data" on Google should present you with a tutorial.

Instagram (instaport.me)

This is not an official Instagram export option, but it will archive an entire account if supplied the target user name and password. Log in with the target Instagram account, select the option to export all photos and videos, and the "Download only" link to download all Instagram content in a compressed file.

International Social Networks

While this book is heavily focused on social networks popular in the United States, they tend to be fairly global with an international presence. This is especially true for Facebook and Twitter. However, there are many social networks which are not popular within the United States that are the primary networks to local residents abroad. This section attempts to identify and explain the most popular foreign networks that may be used by your international targets.

Russia: VK (vk.com)

VK is basically a Russian version of Facebook. You can create a new free account or log in using your existing Facebook credentials. Most search options function without logging into an account. The page at vk.com/people offers advanced search which allows filtering by location, school, age, gender, and interests. Most profiles publicly display a user's birthday, location, and full photo collection. In addition to the publicly visible content, I have found a third-party technique to be beneficial. You should see the user ID number or user name of each VK user within the URL of the profile page. As an example, vk.com/tomcruise would indicate the user name is tomcruise. Placing this detail within a specific URL may reveal a more complete profile of the target. In this example, the following address would be used. This technique has been very hit or miss, but always worth researching.

https://udb.im/vk/user/tomcruise/

The most beneficial tool that I have found associated with VK is **FindFace** (findface.ru). Navigating to this site allows you to upload or link to an image, and the tool will scour VK for any people with matching photos. This is much more than a simple reverse image search. This has been proven to locate additional images of a target from a completely different photo. In one public example, a researcher supplied still captures of pornography actors, and the result was the personal VK profiles of the subjects. If you have a target with an international presence, I highly

recommend an image search here. Note that you must be logged in for this feature to function. Of all of the international social network options, this one has given me the most useful intelligence.

Russia: Odnoklassniki (ok.ru)

Odnoklassniki works similar to most other social media platforms. It is intended to be a way to communicate with friends, as well as an opportunity to network with other people with similar interests. The service is concentrated on classmates and old friends, and translates to "Classmates" in Russian. The official search page is located at ok.ru/search, but you will need to create an account to take full advantage of the options. I have found a targeted site search on Google to be most effective. Searching for Michael Smith would be conducted as follows.

site:ok.ru "michael smith"

The links connect directly to profiles, which can be browsed as normal. These will appear very similar to Facebook profiles. The upper-right portion of a profile will announce the date of the user's last login. Most of the profile details are public, and do not require any type of URL trickery in order to expose the details.

China: QZone (qq.com)

Qzone is typically used as a blogging and diary platform, much like LiveJournal. Most of the loyalty to the platform is due to the popularity of the instant messaging tool "QQ" provided to all users. Unfortunately, this is a one-to-one messaging platform, so opportunities for public search are not present. The search options on qq.com pages provide results similar to Chinese search engines such as Baidu. The searches are not restricted to the social network profiles. I have found the following search on Google or Baidu to work best for English queries. Replace "Michael Smith" with your target's real name or user name.

site:user.qzone.qq.com "michael smith"

China: Renren (renren.com)

Literally translated as "Everyone's Website", Renren is a Chinese remake of Facebook. It is one of the most popular Chinese social networks. Users earn points for activities such as logging in, posting status messages, commenting, and receiving comments. As users earn points, their level on the site increases, which unlocks special emoticons, profile skins, and the ability to go "invisible" and view other users' profiles without their knowledge. The home page does not allow profile search, but browse.renren.com does. Clicking on any profiles from this query will prompt you to sign into an account. However, a targeted site search should eliminate this. The following Google search identified several various pages that contained Michael Smith.

site:renren.com "michael smith"

If the results are too overwhelming, you can use the following structure to filter the content.

site:blog.renren.com "michael smith" (Filter for blog results only)
site:page.renren.com "michael smith" (Filter for profile results only)
site:zhan.renren.com "michael smith" (Filter for news results only)

Latin America: Taringa (taringa.net)

Taringa has a presence in every country in the Spanish-speaking world. Its main markets are Argentina, Spain, Colombia, Chile, Peru, and the US Latino community. The search functionality is fairly straight-forward, but the following URLs may produce more efficient results.

taringa.net/buscar/posts/?q=OSINT (Searches OSINT within post comments)
taringa.net/buscar/comunidades/?q=OSINT (Searches OSINT within community posts)
taringa.net/buscar/shouts/?q=OSINT (Searches OSINT within "Shout" posts)
taringa.net/buscar/imagenes/?q=OSINT (Searches OSINT within images)

These social networks represent only a portion of the available options. If you encounter your target within any of these social networks, you should research additional options for that region. Overall, resort to custom Google searches when the foreign language barriers become an issue.

Everything Else

Readers of my previous editions may notice that this chapter is much shorter than earlier works. Past versions included many screen captures and several services that have since shut down. With respect to the intelligent audience, I refrained from including any images in this chapter. All of these search options are very easy to use and the functions are fairly obvious. The details of this chapter are often the first to become outdated after a new release. Instead of over-explaining all of the new and upcoming social media monitoring options, I leave a list of resources here for you to research. They all have strengths and (many) weaknesses. You may find a perfect solution here, but you are more likely to see more of the same.

KeyHole (keyhole.co)
Board Reader (boardreader.com)
HashAtIt (hashatit.com)
UVRX (uvrx.com/social.html)
Who's Talkin (whostalkin.com)

CHAPTER SEVEN
ONLINE COMMUNITIES

Online communities are very similar to social networks. While social networks cater to a broad audience with many interests, these communities usually relate to a specific service or lifestyle. Many online communities do not get indexed by search engines; therefore, the presence of a target's participation will not always be found through Google or Bing. Any time that a target's interests or hobbies are located through the previous search techniques, you should also seek the online communities that cater to that topic. This can often identify information that is very personal and private to the target. Many people post to these communities without any regard to privacy. Some communities require registration to see the content, which is usually free. Occasionally, a cached version of the pages on the site is available without registering. This chapter will provide methods of infiltrating these communities to maximize the intelligence obtained.

Reddit (reddit.com)

Contrary to previous editions, I now start this chapter with Reddit. This social news aggregation, web content rating, and discussion website went from a place for those "in-the-know" to a resource often cited on mainstream media. More users than ever post, reply to, and read the user-submitted content in the form of either a link or text, each submitted to a specific category known as a Sub-Reddit. Other users then vote the submission up or down, which is used to rank the post and determine its position on the website's pages. The submissions are then discussed on the "comments" page of every entry. The Sub-Reddits cover practically any topic you can imagine. If your target has the slightest interest in the internet, he or she has probably been to Reddit. As of 2018, there were over 900,000 Sub-Reddits and 250 million registered users.

Reddit Search

The official search option on Reddit has been plagued with problems since inception. The search field in the upper right of every page will allow you to query any terms you desire, but the results have not always been optimal. In 2016, I saw the search function improve drastically, and even with some added new features. When typing terms into the search field on any Reddit page, the results will be from all pages, including thousands of Sub-Reddits. While you should understand this option, and even execute target searches from the home page on occasion, we should consider some advanced search options.

We can replicate that standard keyword search within a URL. This is beneficial for bookmarking searches of interest that need checked often. The format for a search about OSINT is as follows.

https://www.reddit.com/search?q=OSINT

The results from such a generic search can be quite overwhelming. With the following URL, we can force Reddit to only deliver results if our search term is within the title of a post, and not simply present anywhere within the comments.

https://www.reddit.com/search?q=title:OSINT

If you know the name of the Sub-Reddit, you can navigate directly with the following structure.

https://www.reddit.com/r/OSINT/

If you locate a user name of interest while searching Reddit, you can load all of that user's posts and comments by clicking on the name. Alternatively, the following URL can be used.

https://www.reddit.com/user/CHRISB

If you have a target website, and you want to know if the URL has ever been posted as a submission link, the following URL will display all results.

https://www.reddit.com/search?q=site:inteltechniques.com

If Reddit is not providing the results you think you should be receiving, you should return to our previous instruction on Google searching. The following query would identify any posts, categories, or users that included the word "surveillance".

site:reddit.com "surveillance"

If you wanted to force Google to restrict its searching to a specific Sub-Reddit, such as OSINT, you would add "/r/osint" after the first portion. If you wanted to restrict the searching to a specific user, you would add "/user/CHRISB" to the end. Specifically, the following would apply.

site:reddit.com/r/osint "surveillance"
site:reddit.com/user/CHRISB "surveillance"

Deleted Content

If you have identified any Reddit content of interest, you should consider checking any online third-party archives. These historic representations of an account will often disclose previously deleted or modified content. It is extremely common for Reddit users to edit or delete a comment entirely, especially if it was controversial. I have investigated numerous Reddit accounts where the evidence I expected to find was not present. First, I always search the standard archive options that were explained in Chapter Three. The following three direct URLs would attempt to display historic versions of a Reddit user's profile. You could replace the Reddit user URL within each of these with a Sub-Reddit address or Reddit post URL.

webcache.googleusercontent.com/search?q=cache:https://www.reddit.com/user/CHRISB
web.archive.org/web/*/https://www.reddit.com/user/CHRISB
archive.fo/https://www.reddit.com/user/CHRISB

You may get lucky with these queries, but the results are often only the tip of the iceberg. These will display the historic view of a Reddit user account at a specific moment in time. While this may provide enough evidence for your investigation, you should also take the next step and identify any user analytics. The following direct URL displays all available metadata for the user CHRISB. This includes an analysis of the user's posts, which identifies the most liked and disliked comments and submissions by the user; his or her activity isolated by year, day of the week, and time of day; and the most common words submitted.

https://snoopsnoo.com/u/CHRISB

This graphical view can be very beneficial in order to obtain a quick review of the target's Reddit usage, but it is only a summary. The best and worst posts may display previously deleted content, but it does not tell the entire story. In order to dig much deeper, we will dig into the Pushshift datasets.

Pushshift (files.pushshift.io)

This huge archive contains over 300GB of data including most publicly posted content on Reddit since 2005. This provides us an amazing collection of most deleted posts. The next time your target wipes out various Reddit posts before you can collect the evidence, Pushshift may reveal the deleted content. This site allows you to download the entire archive, but that may be overkill for most users. Instead, we can take advantage of their robust application programming interface (API). First, let's assume that you are only interested in a specific user that has deleted all content. The following direct URL queries the entire data set for any posts that have been archived by Pushshift.

https://api.pushshift.io/reddit/search/comment/?author=CHRISB

This will display the most recent 25 posts, regardless of whether they are still on Reddit or have been removed. This is a great start, but our target may have thousands of posts. The following URL adds two options at the end to force sorting in ascending format and display 1000 comments within a single page.

https://api.pushshift.io/reddit/search/comment/?author=CHRISB&sort=asc&size=1000

If you are seeking a specific post with unique wording, you can accomplish this with the following URL. This example would identify public and deleted posts mentioning my website.

https://api.pushshift.io/reddit/search/comment/?q=inteltechniques.com

Each of these searches may present too much content and may not be easy to digest. We can filter unwanted content in order to produce less results. The following would repeat our previous search, but only display content from the Sub-Reddit NetSec.

https://api.pushshift.io/reddit/search/comment/?q=inteltechniques.com&subreddit=netsec

If you wanted to limit results to a single user with a timeframe between 5 days prior to your search and 30 days prior to your search, you would navigate directly to the following URL.

https://api.pushshift.io/reddit/search/comment/?author=CHRISB&after=30d&before=5d

Note that all of these searches only identify results that are comments, and not user submissions. A submission is a new topic, and a comment is a post within a specific submission. In order to replicate all of these queries for user submissions, simply replace "comment" in each example to "submission". In order to demonstrate the value of this, consider the following real example.

On December 29, 2017, Reddit user iPhoNewsRO Tweeted "great so my reddit account was hacked and used for scams and now it got deleted. lost 5 years of saved posts". Figure 7.01 displays the current view of the account. On December 30, 2017, I navigated to the following URL, which displayed 100 of the most recently deleted comments. His earliest post was "HMU = hit me up aka send me a DM/PM" to Reddit user "Obey_Kush". Evidence of this interaction is not present anywhere on the live view of Reddit.

https://api.pushshift.io/reddit/search/comment/?author=iPhoNewsRO&sort=asc&size=100

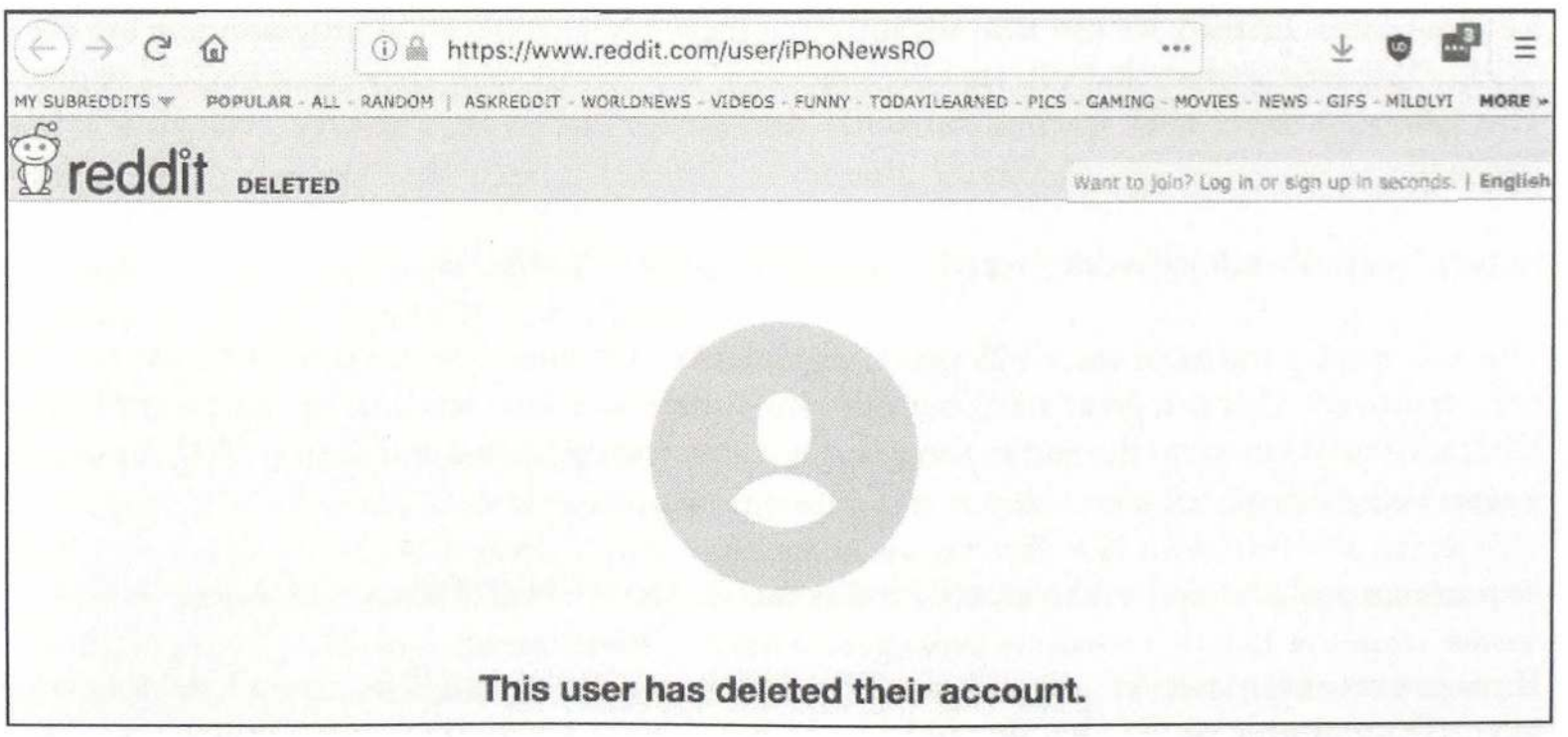


Figure 7.01: A deleted Reddit account that was recovered with Pushshift.

Images

Reddit is well-known for hosting entertaining images and memes. The majority of linked images on Reddit are hosted on a photo-sharing site called Imgur (imgur.com). This can be very beneficial when you are investigating an image post that has been removed from Reddit. If a user posts a photo to Imgur, then links it to a Reddit post, and then deletes the post, the image is still online. You will no longer have a link to the image, and randomly searching Imgur will be unproductive. Instead, we can browse all of the Reddit images on Imgur with a direct URL. The following address will display the current images, in reverse chronological order, associated with the Sub-Reddit titled NetSec. Scrolling will continuously load older images.

https://imgur.com/r/netsec

If you find an image of interest, you should consider a reverse image search. This will be explained in detail later, but you should know now that you have a Reddit-specific reverse image search option call Karma Decay. Assume that you located an image on Imgur at the following URL.

https://imgur.com/r/funny/0DnE1aB

You can navigate to karmadecay.com, supply this address, and immediately see if that image has been posted to any other locations within Reddit. If you wanted to bookmark a direct URL for future checking, you could use the following to obtain the same result.

http://karmadecay.com/imgur.com/r/funny/0DnE1aB

Note that Karma Decay blocks pornographic images. If your investigation involves any adult photos, you will need to use a different service called NSFW Reddit Reverse Image Search (i.rarchives.com). Enter any image into that website and it will search for other copies within Reddit. Optionally, you can submit directly via URL as follows.

http://i.rarchives.com/?url=http://i.imgur.com/mhvSa.jpg

Investigation Sub-Reddits

There are many Sub-Reddits that can provide a unique benefit to an investigator, three of which are outlined here. There are several versions of each of these, but those that I present here have the most history of being helpful. You will find additional options with a bit of searching.

Reddit Bureau of Investigation (reddit.com/r/rbi)

This active community helps other Reddit users solve crimes and other problems. Internet gurus will help find deadbeat parents; computer specialists will aid in tracking stolen devices; and private investigators will assist with investigation techniques. I have used this option several times during

my career. The most successful cases involved hit and run traffic crashes. In 2013, I assisted a northern Illinois police department with the investigation of a fatal car crash. The offender fled the area and an elderly woman died. Three small pieces of the offending vehicle were left at the scene. After posting this information to local media outlets, I submitted it to RBI. Within minutes, several vehicle body shop employees were tracking down the parts and eventually tied them to a specific year and model of a 10-year-old vehicle. This information led to the arrest of the subject. Another victim of an unrelated hit and run traffic crash posted a blurry photo of the suspect vehicle and asked for assistance. Within hours, a Reddit user identified the license plate through digital correction techniques.

Pic Requests (reddit.com/r/picrequests)

A constant frustration in my work is blurry, out of focus, or grainy digital images. Commonly, I will receive surveillance photos that are too dark or light to identify anything of value in the image. Occasionally, I will find images on social networks that could be beneficial if they were just a touch clearer. Pic Requests saves the day. This Sub-Reddit consists of digital photo experts that can perform Photoshop magic on practically any image. Many Reddit users will request old photos colorized, torn photos digitally repaired, or unwanted subjects removed from an image. I have uploaded several surveillance images to this group with a request for assistance. The users have been incredibly helpful by identifying digits in blurred license plates and turning dark surveillance footage into useful evidence.

What Is This Thing? (reddit.com/r/whatisthisthing)

I am consistently amazed at the results from this Sub-Reddit. What Is This Thing is a place that you can post a digital photo of practically anything, and someone will know exactly what it is while providing detailed and cited additional information. Many users post images of old antiques and intricate items hoping to identify something valuable in their collection. I use it to identify tattoo meanings, graffiti, suspicious items mailed to politicians, vehicle parts, and just about anything else that is presented to me during my investigations.

Real World Application: In 2012, I was asked to assist with a death investigation of a "Jane Doe". I submitted a sanitized version of a tattoo on her back that appeared to be Chinese symbols. Within five minutes, a Reddit user identified the symbols, their meaning, and references to the region of China that would probably be related to my investigation. A reverse image search of his examples led to information about a human trafficking ring with which the victim was associated. This all occurred over a period of one hour.

If you plan to use these techniques on Reddit, please consider a few things. You should create a free account now and hold on to it. Creating a new account and asking for help minutes later can be viewed as rude. I like to use accounts that appear to have been established a long time ago. If you are visible as an active member of Reddit with a history of comments, this might encourage other active members to assist you. You should never be demanding in your requests. Remember,

these people are volunteering to help you. Many of them possess a skill set that cannot be found elsewhere. I also never upload any content that is not already publicly available. If digital images were released to the press, I have no problem releasing them to Reddit. If my target image is already on a public social network, I see no reason it cannot be linked through Reddit.

While Reddit seems to get most of the attention in this type of community, there are alternative options that are growing rapidly. These include Voat, 4chan, Hacker News, and others. I will briefly discuss the most common search options, which can be replicated with my Custom Communities Search Tool that is explained later.

Voat (voat.co)

When launched in 2014, Voat appeared to be one of many Reddit clones that was surfacing. In June of 2015, it gained a lot of steam as a solid Reddit competitor due to some backlash in the community. At that time, Reddit had just banned several Sub-Reddits that were known to contain harassing content. Many of the hundreds-of-thousands of subscribers that felt betrayed by Reddit fled to Voat as a replacement. It is now known as the home for former Reddit early adopters. While it is very similar to Reddit, you will immediately notice that there is absolutely no search option on the site. Instead, we must rely on searchvoat.co. The search fields at the top of this site are self-explanatory. For those that prefer direct URL access, or those that wish to save specific searches for future use, the following apply. These are very similar to Reddit. Notice the inclusion of "&b=on" (searches body of posts) and "&nsfw=on" (searches adult content).

Text Search: https://searchvoat.co/?t=OSINT&b=on&nsfw=on
User Search: https://searchvoat.co/?u=CHRISB&b=on&nsfw=on
Domain Search: https://searchvoat.co/?d=inteltechniques.com&b=on&nsfw=on
Subverse (Sub-Reddit) Search: https://searchvoat.co/?s=OSINT&b=on&nsfw=on
Subverse (Sub-Reddit): Page: https://voat.co/v/OSINT
Google Search: https://www.google.com/search?q=site:voat.co+OSINT

4chan (4chan.org)

4chan is a mess. It is an image-board website and users generally post anonymously, with the most recent posts appearing above the rest. 4chan is split into various boards with their own specific content and guidelines, modelled from Japanese image-boards. The site has been linked to internet subcultures and activism groups, most notably Anonymous. The site's "Random" board, also known as "/b/", was the first board to be created, and is the one that receives the most traffic. This site is full of bullying, pornography, threats, and general illicit behavior. It has also been the focus of numerous investigations. Similar to Voat, there is no search feature. In this scenario, we will use 4chansearch.com. The following examples are direct URLs that take advantage of this third-party search option, each using "OSINT" as a search term.

Active Search: http://4chansearch.com/?q=OSINT&s=4
Archives Search: http://4chansearch.com/?q=OSINT&s=7
Archives Alternative: https://archive.4plebs.org/_/search/text/OSINT/order/asc/
Google Search: https://www.google.com/search?q=site:4chan.org%20OSINT

Hacker News (news.ycombinator.com)

While this site is targeted toward a tech-savvy community, general discussion topics are followed by millions of users daily. Fortunately, we have a lot of control with searching specific posts, keywords, users, and favorites. The following searches locate data based on a keyword (OSINT) and user (CHRISB).

Text Search: https://hn.algolia.com/?query=OSINT&type=all
User name Search: https://news.ycombinator.com/user?id=CHRISB
User Posts: https://news.ycombinator.com/submitted?id=CHRISB
User Comments: https://news.ycombinator.com/threads?id=CHRISB
User Favorites: https://news.ycombinator.com/favorites?id=CHRISB
Google Search: https://www.google.com/search?q=site:news.ycombinator.com+OSINT

A few Reddit alternatives that are currently small but show promise of becoming relevant include the following, with the best search option for each.

Raddle: https://www.google.com/search?q=site:raddle.me+OSINT
Steemit: https://www.google.com/search?q=site:steemit.com+OSINT
Hubski: https://hubski.com/search?q=OSINT

Meetup (meetup.com)

Meetup consists of users and groups, with all communication related to events where people actually meet in real life. Each user creates a profile that includes the person's interests, photos, and user name. A group is created by one or more users and is focused on a general interest in a specific location. An example would be the "Houston Dog Park Lovers", which is a Houston-based group of dog owners that meet at dog parks. Each group will post events that the members can attend. The majority of the events posted on Meetup are visible to the public and can be attended by anyone. Some groups choose to mark the details of the event as private and you must be a member to see the location. Membership is free and personal information is not required.

You can search Meetup by interest or location on practically any page. Once you locate a group page, you can browse the members of the group. This group page will also identify any future and past events sponsored by the group. These past events will identify the users that attended the event as well as feedback about the event. This site no longer offers the option to search by user name. In order to do this, you will need to use a search engine as described in a moment. A user profile will often include links to social networks and messages from associates on the

website. Additionally, these profiles will identify any future Meetup events that the user plans on attending. Because of this, the site has been used in the past by civil process servers, detectives, and the news media to locate people that had been avoiding them. The following Google search structures have been most helpful in my experience.

Name Search (John Morrison): site:meetup.com inurl:member john morrison
Event Search (Protest): site:meetup.com inurl:events Protest
Post Search: site:meetup.com inurl:discussions Protest
Location Search (Zip-62025): meetup.com/find/events/?radius=5&userFreeform=62025
Google Keyword Search (OSINT): site:meetup.com OSINT

Dating Websites

When investigating cheating spouses, background information, personal character, or applicant details, dating sites can lead to interesting evidence. The presence of a dating profile does not mean anything by itself. Millions of people successfully use these services to find mates. When a target's profile is located, it will usually lead to personal information that cannot be found anywhere else. While many people may restrict personal details on social networks such as Facebook, they tend to let down their guard on these intimate dating websites. In my experience, the following will apply to practically every dating website.

- You must have an account to browse profiles, which is usually free
- You must have a premium (paid) account to contact anyone
- If a target uses one dating service, he or she likely uses others

Instead of explaining each of the dating services, I will focus on methodology of searching all of them. While each website is unique and possesses a specific way of searching, they are all very similar. Overall, there are three standard search techniques that I have found useful, and they are each identified below.

User name: Every dating website requires a user name to be associated with the profile, and this data is searchable. Surprisingly, most users choose a user name that has been used somewhere else. I have seen many dating profiles that hide a person's real name and location, but possess the same user name as a Twitter account. The Twitter account then identifies name, location, and friends. Additional user name search tools will be presented in Chapter Nine.

Text Search: This is a technique that is often overlooked. Most dating network profiles include an area where the users can describe themselves in their own words. This freeform area often includes misspellings and obvious grammatical errors. These can be searched to identify additional dating networks since many users simply copy and paste their biography from one site to another. In 2013, I was teaching an OSINT course in Canada. During a break, one of the attendees asked for assistance with a sexual assault case that involved the dating website Plenty Of Fish. The unknown suspect would meet women through the online service and assault them.

All of the information on his profile was fake, and the photos were of poor quality and unhelpful. Together, we copied and pasted each sentence that he had written in his bio for the profile. Eventually, we found one that was very unique and grammatically worded poorly. A quoted Google search of this sentence provided only one result. It was the real profile of the suspect on Match.com, under his real name, that contained the same sentence describing himself. The high-quality photos on this legitimate page were used to verify that he was the suspect. An arrest was made within 24 hours.

Photo Search: In Chapter Fourteen, I explain how to conduct reverse-image searching across multiple websites. This technique can compare an image that you find on a dating network with images across all social networks, identifying any matches. This will often convert an "anonymous" dating profile into a fully-identifiable social network page. This applies to any dating networks, and photos will be your most reliable way of identifying your target.

The list of popular dating websites grows monthly. The following are the current most popular services. I also maintain a custom search page that queries Google for any user name or keyword at https://inteltechniques.com/OSINT/dating.networks.html.

Match (match.com)
Plenty of Fish (plentyoffish.com)
eHarmony (eharmony.com)
OK Cupid (okcupid.com)
Christian Mingle (christianmingle.com)
Ashley Madison (ashleymadison.com)
Adult Friend Finder (adultfriendfinder.com)
Farmers Only (farmersonly.com)
Elite Singles (elitesingles.com)
Zoosk (zoosk.com)
Friendfinder-X (friendfinder-x.com)
Badoo (badoo.com)
Tinder (tinder.com)

A section about online dating would not be complete without a reference to Tinder. While this service was natively available only through a mobile app, they have recently begun allowing account access via their website. There are hurdles with this, but we can work through all of them. The simplest explanation of Tinder is that it connects you with people in your immediate area who are also using the service. Some call it a dating app, some refer to it as a "hook-up" app. Either way, it is probably the most popular dating service available today. Note that access to Tinder through their app will be explained in Chapter Twenty-One. In order to access Tinder from your web browser, several things must be perfectly aligned in order to prevent account blocking. Tinder gets bombarded with fraudulent accounts, and their radar for investigative use is very sensitive. The following instructions assume that you do not have an existing Tinder account.

- Open the Chrome browser and search for Manual Geolocation Chrome
- Install the Manual Geolocation Chrome extension
- Click the new icon in the toolbar and set your location as desired
- Navigate to tinder.com through the Chrome browser

- Click the Login button and choose Log In With Phone Number
- Supply a Google Voice number
- Confirm the text message received on Google Voice number
- Close Tinder tab and reopen Tinder.com in Chrome
- Click the Login button and choose Log In With Phone Number
- Supply the same Google Voice number
- Confirm the text message received on Google Voice number
- Complete registration with alias name and cropped public domain photo

These instructions may seem a bit weird and redundant, but I will explain my process. I have found that successfully spoofing GPS is more reliable in Chrome than Firefox. I have also witnessed Firefox block elements required by Tinder, likely due to default privacy settings. While you might be able to make Firefox work with Tinder, this Chrome process simply worked better for me. There are many GPS spoofing extensions available for Chrome and Firefox, and most should work. I found the Manual Geolocation option especially successful as both it and Tinder use HTML5 standards for location data. Surprisingly, I had great luck with using Google Voice numbers for web verification. These never work for me on the app, so this is a great way to create the accounts. When I tried other telephony services, such as Sudo, Burner, and Blur, I was denied a new account. After initially confirming my Google Voice number, Tinder did not allow me to alter my profile settings. I had to close the tab and login a second time in order to be prompted by Tinder to finish creating my account. Your mileage may vary. During registration, you must upload a photo. I recommend royalty-free images such as those found on pikwizard.com. They are free, require no licensing, and do not compromise you or anyone else. I searched the word "group" and found numerous appropriate images.

After you have successfully completed the account creation process, the sky is the limit. Tinder will detect the location that you specified in the extension, and start to identify people in the area. Click "My Profile" in the upper left to choose your desired distance, age range, and other details. I always recommend lowering the distance in order to have more control of the target area. The lowest setting has been most appropriate for my investigations. If you change your location in the extension, close Chrome and reopen Tinder to apply the new setting to your account. In order to login to the account later, simply use the Phone Number option and repeat the message process. You should be able to use your new account within the emulator discussed later.

Real World Application: For years, I have experienced difficulty creating new accounts on Tinder from within an emulator, as explained in Chapter Twenty-One. In late 2017, I needed to create five new Tinder accounts for an active investigation into a sex trafficking incident where a pimp was selling a 16-year-old girl on Backpage. After removing all public posts, an informant notified us that the pimp had moved to Tinder, as it was more "private". I created five accounts with five Google Voice numbers from five VPN IP addresses, all within ten minutes (I tried within five minutes, but failed). After creation, I set my GPS within each account to one mile from the motels where the pimp often worked. I did this through multiple copies of Chrome

inside virtual machines such as Buscador. I quickly located two images of what appeared to be young girls in a shady motel room. I "swiped right" to indicate interest, and was immediately contacted by the pimp pretending to be one of the girls. We agreed on a price and he disclosed the room number. The local agency working the case began immediate surveillance while a search warrant was obtained. While waiting, they arrested two men before the "date" could begin, and the pimp after a search warrant was obtained for the room.

Ethnicity Specific Communities

There are several social/dating websites that focus on one specific race. While these communities do not prohibit members of another race from joining, the great majority of the members are of a single race. Black Planet (blackplanet.com) has a large African-American presence, MiGente (migente.com) has a large Latino-American presence, and Asian Avenue (asianave.com) has a large Asian-American membership. Each of these has extensive search capabilities. The easiest option is to provide a user name which may identify a target's profile.

Forums: Board Reader (boardreader.com)

Online forums provide a unique place for discussion about any topic. If you can think of the subject, an entire site full of people is probably hosting a discussion about the topic. These are usually referred to as user forums. Sometimes, these sites are excluded from being indexed by search engines. This can make locating them difficult. A new wave of forum search sites fills this void. Each of the following sites will search the forum communities, message boards, discussion threads, and other general interest groups that post messages back and forth. Board Reader offers an advanced search which allows you to choose keywords, language, data range, and specific domain. If you have trouble filtering results on other forum search sites, this can be useful.

Forums: Craigslist (craigslist.org)

This forum is categorized by topic instead of location, but location filtering is supported. These forums are not indexed by most search engines, so a manual search is the only way to see the content. In order to search these areas, you must create a free user account. As usual, you can use fictitious information in your profile. After logging in, you can search by keyword on this main page, but not by screen name. This option will identify posts matching the search terms within any of the topics.

The "Handle" option will search by user name but can only be seen by clicking on any topic. I entered the "Yoga" room which displays the additional "Handle" search option. This will identify the posts of an individual user. I have found this "handle" option useful to an investigator. As a general rule, most people will use the same user name across several sites. Craigslist is no exception. If you have identified a user name of a target, a search on the Craigslist forums is worth a look. Although you will not get a result every time you search, the commentary is usually colorful when you do. When you locate a user name of a target on the Craigslist forums, searching

that user name will provide an abundance of information within the user's profile page. These often display a forwarding email address, date joined, photograph, and up to 50 posts made during the past 31 days. These can provide great intelligence on the target.

Forums: The Hood Up (thehoodup.com)

This interesting website has a subtitle of "Where American Hoods Connect". The index page has categories for open discussion, the East Coast, the West Coast, the Midwest, and "Down South". The premise of the site is a place for "gangsters" and "hoods" to communicate with each other in an uncensored environment. A quick look into the Midwest group displayed numerous conversations about criminal activity, gangs, and violence. Each page has a search option for any keyword. A search of "Chicago Gaylords" resulted in 799 posts discussing everything from the history of the gang to debates about current and future plans. Any of this would be great intelligence for the gang units of the Chicago Police Department.

Online Newspaper Comments

Practically every newspaper now has some sort of online presence. Most digital editions allow readers to leave comments about individual articles. These comments can usually be seen at the bottom of each web page. While the value that these comments add to the newsworthiness of each piece of news is debatable, the content can be important to an investigation. In years past, most newspapers hosted their own digital comment delivery system within their website. This often resulted in a large headache while trying to maintain order, prevent feuds between readers, and delete direct threats. Today, most news websites use third party services to host these comments. The most popular are Facebook and Disqus. When Facebook is utilized, most people use their real names and behave better than when using only a user name on Disqus. Any complaints about the comment activity can be referred to Facebook since they technically store the content. Searching Facebook comments can be conducted through the technique explained in Chapter Three.

In order to search for content within the Disqus comment system, you can conduct a custom Google search. First, it is important to understand how the Disqus system is recognized by Google. There is an option to log into a Disqus account and you can "upvote" or "downvote" each comment to show your approval. The words visible on this page that were provided by Disqus are important for the search. The word "comments" will be visible on every Disqus provided environment and there will also be a link to disqus.com. Therefore, the following search on Google should provide any websites that have the Disqus comment delivery system and also have a reference to OSINT.

"osint" "disqus" "comments"

This may produce some non-Disqus results that happen to possess all three words, but those should be rare. This will also identify many pages that do not contain any comments whatsoever.

In order to only receive results that actually have comments, alter your search to the following.

"osint" "disqus" "1..999 comments"

This instructs Google to only display results that contain the keywords "OSINT" and "Disqus" and also contain the exact phrase of any number between 1 and 999 followed immediately by the term "comments". This would provide results that contain any number of comments with the exception of "0" or over "1000". The "1..999" portion is the Google range operator that will display any number within the specified range.

Online Prostitution

Craigslist was once used by many prostitutes nationwide as an avenue to meeting "Johns". Likewise, many people used the site to locate a prostitute. In 2009, Illinois Attorney General Lisa Madigan convinced Craigslist to remove the "Erotic Services" section that hosted these posts announcing illegal activity. Today, it is difficult to find a post offering prostitution on Craigslist. Unfortunately, this does not mean that the prostitutes and their clients simply stopped the illegal behavior. Instead, they found new resources. There are many sites online that aid in prostitution and human trafficking. A few of the big players are listed here.

Backpage (backpage.com)

When Craigslist turned off the "Erotic Services" section, the traffic immediately went to Backpage. Prior to 2016, after selecting your location on this site, you were presented with a main page of categories including the "Adult" area. This included subsections of Escorts, Body Rubs, Stripper & Strip Clubs, Dom & Fetish, Trans-Sexuals, Male Escorts, and Adult Jobs. It should be no surprise that the posts inside of these sections included blatant ads for prostitution. This entire section was removed. Unsurprisingly, the sex workers simply moved the posts into the "Dating" section. Almost all posts in this section include photos, many of them nude, and prices for various services. Cellular telephone numbers are common, but are usually spelled out in text. Instead of typing an area code of "314", the user may type "three one four". Because of this, an investigator must get creative with the searches. A search field is at the top of every page and allows for any keyword search. Several juvenile prostitution cases have started at Backpage.

Escort Review Websites

These types of services may be difficult for some readers to understand. I was also surprised when I first found them. This is where prostitution clients communicate with each other and leave reviews of their experiences with the prostitutes. These "Johns" document the slightest details of the experience including price, cleanliness, and accuracy of the photograph in the ad. Furthermore, this is the first location that will announce an undercover operation by the police. This is important for law enforcement, as this could create an officer safety issue. It is also how the police can determine when a new name or photo should be used in future covert ads. Another

purpose for this data is to create documentation of the reviews of an arrested prostitute. This can prove valuable in court for the prosecution of offenses. There are several of these services, and every metropolitan area will have a preferred website by the customers. A Google search of "Escort reviews Anaheim" will get you to the popular options. Of course, replace Anaheim with your city of interest. The following techniques can be useful when you identify the service relevant to your investigation.

The Erotic Review (theeroticreview.com)

If you do not know of any individual services that prostitution clients are using in your area, The Erotic Review is a safe bet. Practically every metropolitan area has a presence here. Much of this site will not be available unless you join as a premium member. However, there should be plenty of visible free content for basic investigations. Most of the posts are unsuitable for this book.

Rate That Provider (ratethatprovider.com)

Rate That Provider, and many others, requires an account to access the forums. This can be covertly created, or you can often use a search engine to get around the limitation. A search on Google of "site:ratethatprovider.com chicago", without the quotes, identifies 32,500 results of forum posts. Usually, this link will forward you to a login page. Credentials must be provided before you can see the content. Accessing the Google Cache of each link will likely bypass this demand and display the content information without membership. Practically every post on this forum can be accessed with this method.

Escort Ads (escortads.xxx)

This adult website sponsored by pornography ads allows you to enter the cellular telephone number of a suspected sex provider. The results will include all of the profiles that have been associated with the number. It will also combine these results and display a convenient summary of the cities where ads were posted, ages used in the ads, and every photo posted. It will even remove the duplicate photos.

City Vibe (cityvibe.com)

This site offers a search for "escorts" after choosing a geographical location. This will present a page full of posts from local women offering various services for money. Many include a nude photograph, details of the service provided, breakdown of the fees, and whether they will travel to a location (outcall) or demand a client come to them (incall). Several also offer a personal website link with even further information. It is also common for the user to include a cellular telephone number to be used for contact. The search function on this site is fairly weak. Searching a user's name will usually present results, but searching a full telephone number often returns no results. To correct this, only search the last four digits of a telephone number to get a result. Many times, doing this will discover many women using the same contact number, which often

connects to a pimp. Law enforcement takes advantage of this site to set up prostitution stings.

Real World Application: While participating in an FBI Operation, I focused on locating juvenile prostitutes and women forced into the sex industry by pimps. One easy way to determine if a sex worker was traveling extensively was to search her number through the Escort Ads website. If it returned numerous cities with postings, that was a strong indication that she was a full-time sex worker and was likely not traveling alone. Every contact that we made with traveling prostitutes resulted in the identification of the pimps that transported them to the stings.

Craigslist (craigslist.org)

Craigslist is one big online classified ad for every area of the world. The site can ease the pain of finding an apartment, provide numerous options for buying a vehicle locally, or assist in locating just about any item or service that you can imagine that is within driving distance of your home. It is also a landing spot for stolen goods, illegal services, and illicit affairs. While Craigslist offers a search option, the results are limited to active posts only. You can also only search within one category at a time. You can browse through the posts individually, but this will be overwhelming.

Government and private investigators have found much success in locating stolen goods within this site. To start, you must find the Craigslist site for your area. Often, simply visiting craigslist.org will direct you to the landing page for your geographical area. If this does not happen, navigate through your country, then your state, then your metropolitan area to see listings around you. If the theft occurred recently, a live search in the "for sale" section may produce results. I do not recommend searching from the main page, as there are no advanced options. Instead, click on any section title. For example, clicking on the "for sale" section will take us to that area. The top of the page will have a search field that will search all of the categories in this section. Additionally, we can filter by price range, posts that contain images, or terms that only appear in the title of the post.

Craigslist recently added new features that allow you to view results by list view, gallery view, or map view. These locations will only refer to the city of the item, and not exact GPS location. The gallery view can be used as a "photo lineup" to identify a stolen item. The map view can be beneficial when only looking for items within surrounding areas. Four new options on the upper right of every result page allow you to sort the items by newest listings (default), relevance, lowest price, and highest price. Most pages with items for sale will also allow you to filter the results so that only items being sold by individuals are listed. This would eliminate businesses and dealers. The default is to show both, and I recommend leaving that unless you are overwhelmed by the number of results.

If a thief sells the item on Craigslist, he or she will usually delete the post after the transaction is complete. If the post is deleted, it will not be listed in the results of a search on Craigslist. This is where Google and Bing come in. Both Google and Bing collect information from Craigslist posts to include in their search results. This collection can never be complete, but a large archive of

posts is available. Searching Google or Bing with "site:craigslist.org" (without quotes) will search through archived posts on Craigslist that are both active and removed. Similar to the previous example, you can search "site:craigslist.org laptop Edwardsville" without the quotes. This search produced 572 results that match these criteria on Google. These include the current posts that were available with the live search on craigslist.org as well as posts that have been recently deleted from Craigslist. If you wanted to focus only on a specific regional area of Craigslist, changing the search to "site:stlouis.craigslist.org laptop Edwardsville" would filter results. This example would only show listings from the St. Louis section of Craigslist. You can use any region in your custom searches.

The results that are still current will link to the actual post and display all content of the post. If a search result links to a post that has been deleted from Craigslist, a standard "page not found" error will be returned. You can still get additional information from this deleted post by looking through the text supplied on this search page. The brief description will often disclose an email address or telephone number. Some listings may have a cached view, but lately this has been rare. In a scenario where thousands of search results are presented by Google or Bing, you can add search terms to filter to a more manageable amount of posts. Adding the make or model number of the product may quickly identify the stolen property.

You can also search by terms other than the product of interest. Many people that use Craigslist do not want to communicate through email sent from the website. Most users will include a telephone number in the post as a preferred method of communication. The overwhelming majority of these telephone numbers belong to the cellular telephone of the user submitting the post. This can be a huge piece of intelligence for an investigator attempting to identify a person associated with a telephone number. It is common that a criminal will purchase a cellular telephone with cash and add minutes to it as needed. This makes it difficult for someone to identify the criminal from the phone number. Ironically, the same criminal will post the telephone number as well as a name on a public internet site for the world to see. Sometimes, a person will post both a cellular and a landline telephone number on the same post. This allows an investigator to associate these two numbers, and a quick internet search should identify the owner of the landline telephone number.

Another way to search Craigslist posts is to identify screen names within the post. Craigslist discourages inserting a screen name or email address within a post; however, most people have figured out how to bypass this limitation. Instead of someone typing their email address within their posts, they will insert spaces between the first portion of the email address (user name) and the second portion of the email address (domain name). For example, instead of the user typing their email address as JohnDoe911@gmail.com, he or she may identify the account as "JohnDoe911 at gmail com". This would be enough to prevent Craigslist's servers from identifying the text as an email address and prohibiting the post. Fortunately for the investigator, this information is indexed by Craigslist and other search engines to be retrieved.

You can search any keyword in either the official Craigslist site or on Google and Bing using the

"site" operator. In my experience, Bing offers more results of archived Craigslist posts than Google. If you do not have success with Bing, Google should be searched as well. Many private investigators find the "personals" section of interest. The "Casual encounters" area is well known for extramarital affairs. If you want to only search all live Craigslist posts, regardless of which geographical area it exists, you can use sites such as totalcraigsearch.com, adhuntr.com, and searchalljunk.com.

Craigslist has a few advanced search operators that may be of interest. It supports a phrase search with quotation marks such as "low miles". It accepts the hyphen (-) operator to exclude terms such as honda black -red. This search finds postings that have 'honda' and 'black' but not 'red'. A pipe symbol (|) provides "OR" searches such as honda | toyota. This search finds postings that have 'honda' or 'toyota' (or both). You can group terms together in parentheses when queries are complicated. A search of red (toyota | honda) -2000 -2001 finds listings that have 'red' and either 'honda' or 'toyota' (or both) but do not have 2000 or 2001. Wildcards are accepted as follows.

hond* civ* (match "honda civic", "honda civil", etc)
wood floo* (matches "wood floors", "wood flooring", etc)
iphone* (matches "iphone", "iphones", "iphone5", etc)

Craigslist's email alert feature has made third party tools for this purpose unnecessary. After logging into your account, you can customize alerts to send an email to you when specific search terms are located.

Real World Application: Many thieves will turn to the internet to unload stolen items. While eBay requires banking information or a credit card to use their services, most thieves prefer Craigslist's offer of anonymity. My local police department successfully located a valuable stolen brass instrument this way and set up a sting to arrest the thief. Often, the thief will be willing to bring the item to you in order to get some quick cash. Another tip that has helped me during investigations is to look for similar backgrounds. When I had a group of gang members stealing iPhones from vehicles and pockets, they would sell them right away on Craigslist. Since there were hundreds of legitimate iPhones listed, identifying the stolen units can be difficult. By looking for similarities in the backgrounds, I could filter the list into interesting candidates. Finding unique backgrounds, such as tables or flooring, within several posts can be suspicious. Additionally, I have found posts that include "hurry", "must sell today", and "I will come to you" to be indicators of illegal activity.

eBay (ebay.com)

eBay is an online auction site. Since the site requires a user's financial information or valid credit card to post items for sale, many thieves have moved to Craigslist to unload stolen goods. eBay offers an advanced search that will allow filters that limit to auctions from a specific location, or specified distance from the location. On any search page, there is an "Advanced" button that will display new options. Of these options, there is a category titled "show results". The last option

in this category is titled "items near me". Here, you can select a zip code and filter results to a minimum of 10 miles from the zip code selected. This will now allow you to search for any item and the results will all be from sellers near a specific zip code. This location option will remain active as you search for different keywords. These searches will only search current auctions that have not expired. In order to search past auctions, select the "Completed listings" option under the category of "Search including". If you want to conduct your searches directly from a URL, or if you want to bookmark queries that will be repeated often, use the following structure. Replace TERMS with your search keywords and USER with your target's user name.

Keyword: ebay.com/dsc/i.html?&LH_TitleDesc=1&_nkw=TERMS
Sold: ebay.com/sch/i.html?_from=R40&_nkw=TERMS&LH_Sold=1&LH_Complete=1
Completed: https://www.ebay.com/sch/i.html?_from=R40&_nkw=TERMS&LH_Complete=1
User name: https://www.ebay.com/usr/USER
User Feedback: https://feedback.ebay.com/ws/eBayISAPI.dll?ViewFeedback2&userid=USER
User Items: https://www.ebay.com/sch/USER/m.html
User Search: http://www.ebay.com/sch/ebayadvsearch/?_ec=104&_sofindtype=25&_userid=USER
User Followers: https://www.ebay.com/usr/USER/followers#followers
User Following: https://www.ebay.com/usr/USER/all-follows?priflwtype=people#people

Flippity (flippity.com)

An alternative to the location feature on the official eBay site is Flippity. This site performs the same function as mentioned previously, but with less work on the user's part. The results of your search will appear on a map with the ability to minimize and expand the radius as desired. This is a quick way to monitor any type of items being sold in a specific community.

GoofBid (goofbid.com)

Not everyone uses spellcheck. Some people, especially criminals, will rush to list an item to sell without ensuring that the spelling and grammar are correct. You could conduct numerous searches using various misspelled words, or you can use GoofBid. This site will take your correctly spelled keyword search and attempt the same search with the most commonly misspelled variations of the search terms. Once this helped me identify a thief selling a "saxaphone". Another alternative to this service is **Fat Fingers** (fatfingers.com)

Search Tempest (searchtempest.com)

If you find yourself searching multiple geographical areas of Craigslist and eBay, you may desire an automated solution. Search Tempest will allow you to specify the location and perimeter for your search. It will fetch items from Craigslist, eBay, and Amazon. You can specify keywords in order to narrow your search to a specific area. Advanced features allow search of items listed within the previous 24 hours, reduction of duplicates, and filtering by categories. While I encourage the use of these types of services, I always warn people about becoming too reliant on them. These tools could disappear. It is good to understand the manual way of obtaining data.

OfferUp (offerupnow.com)

This service is steadily stealing the audience currently dominated by Craigslist. OfferUp claims to be the simplest way to buy and sell products locally. A search on their main page allows you to specify a keyword and location. The results identify the usual information including item description and approximate location. OfferUp follows the eBay model of including the seller's user name and rating. The unique option with OfferUp is the ability to locate the actual GPS coordinates associated with a post instead of a vague city and state. This information is not obvious, but can be quickly obtained. While on any post, right-click and choose to view the page source. Inside this new tab of text should be two properties titled offerup:location:latitude and offerup:location:longitude. You can search for these in your browser by pressing ctrl-f (Windows) or command-f (Mac). Next to these fields should display GPS coordinates. In my experience, these precise identifiers will either identify the exact location of the target, or a location in the neighborhood of the suspect. I would never rely on this all the time, but I have had great success getting close to my targets through this technique.

Amazon (amazon.com)

Amazon is the largest online retailer. Users flock to the site to make purchases of anything imaginable. After the receipt of the items ordered, Amazon often generates an email requesting the user to rate the items. This review can only be created if the user is logged into an account. This review is now associated with the user in the user profile. An overwhelming number of users create these product reviews and provide their real information on the profile for their Amazon account. While Amazon does not have an area to search for this information by user name, you can do it with a search engine. A search on Google of site:amazon.com followed by any target name links to an Amazon profile and several item reviews. The first link displays the user profile including photo, location, and the user's review of products purchased.

This technique of using Google or Bing to search for profiles on websites that do not allow such a search can be applied practically everywhere. Many sites discourage the searching of profiles, but a search on Google such as "site:targetwebsite.com John Doe" would provide links to content matching the criteria. The difficulty arises in locating all of the sites where a person may have a profile. By now, you can search the major communities, but it is difficult to keep up with all of the lesser known networks. This is where the user name search engines mentioned earlier assist.

Amazon does offer a native Wish List search option. If you navigate to their internal page at amazon.com/gp/registry/search, you can type any name or email address to search for that target's wish list, baby registry, or wedding registry. I have found this type of information useful during criminal interrogations. While these details do not insinuate any criminal activity, telling a suspect that I know this information can be interesting. I once had a suspect convinced that I had no evidence against him in reference to a child pornography investigation. When I "slipped" and told him I knew the books that he wanted for Christmas two years prior, a look of fear

replaced his arrogant attitude. He began questioning his lack of confidence in my investigation. He later figured out I simply looked at his Amazon wish list, and accused me of "playing dirty".

FakeSpot (https://www.fakespot.com/)

There is an abundance of fake reviews on Amazon, which can make it difficult to determine which reviews accurately describe a product and which are provided by employees associated with the seller. FakeSpot attempts to identify products that are likely misrepresented by the review community. During a search for a remote-controlled drone, I found that the Amazon "Best Seller" possesses over 53% fake reviews, and top reviewers "tri nguyen" and "EUN SUN LEE" appear to be automated reviewers based on other products. This service also supports analysis of reviewers on Yelp and Trip Advisor.

Pinterest (pinterest.com)

Pinterest is an online "pinboard" where users can share photos, links, and content located anywhere on the internet. It is a way to rebroadcast items of interest to a user. People that follow that user on Pinterest can keep updated on things that the user is searching and reading. The search feature on the main website is useful for keyword searches only. It will search for any term and identify posts that include those words within the description. A search of my last name displayed several photos of people. Clicking each of these links will present the full-page view of the photo and any associated comments. This page will also identify the full name of the person that uploaded the content and the original online source. Clicking on the full name of the user will open the user profile which should include all "pinned" content. Unfortunately, you cannot search a person's full name or user name on Pinterest and receive a link to their profile page. To do this, you must use Google. The following direct search URLs will identify the user names (BILL) and keywords (CRAFTS) present on Pinterest.

User name: https://www.pinterest.com/BILL/
User Pins: https://www.pinterest.com/BILL/pins
User Boards: https://www.pinterest.com/BILL/boards
User Followers: https://www.pinterest.com/BILL/followers/
User Following: https://www.pinterest.com/BILL/following
Pins Search: https://www.pinterest.com/search/pins/?q=CRAFTS
Boards Search: https://www.pinterest.com/search/boards/?q=CRAFTS
Google Search: https://www.google.com/search?q=site:pinterest.com+CRAFTS

Stumble Upon (stumbleupon.com)

A service similar to Pinterest is Stumble Upon. It is a way for people to share their favorite links with their friends. Searching directly within the service is difficult, if not impossible. The home page requires you to be a registered user to proceed. However, using a "site" search in Google will connect you directly to any pages associated with the search. If you know the target's user

name, a direct link will present their entire profile which may also disclose associates. The following URLs are most appropriate.

Google: site:stumbleupon.com "KEYWORD"
User name: https://www.stumbleupon.com/stumbler/BILL
Followers: https://www.stumbleupon.com/stumbler/BILL/followers
Following: https://www.stumbleupon.com/stumbler/BILL/following

Topix (topix.com)

Topix began as a news aggregator which categorized news stories by topic and geography. It quickly became a playground for internet trolls and colorful commentary. While the service is now focusing on entertainment news and content creation, the local communities are still very active. In this area, you will find reviews and complaints about businesses, individual defamation, and the occasional threat of bodily harm. Navigating to the home page will only present celebrity gossip, but navigating to the local forums will present much more value to investigators. The following pages present the most lucrative areas for my investigations.

Local News & Comments: http://www.topix.com/pick-local
General Forums: http://www.topix.com/forum/recent
User Forums: http://www.topix.com/userforum/recent
Location Selection: http://www.topix.com/city

You will likely see various news articles, each of which might have commentary at the bottom of the article. If you want to see the most recent comments on any post within the target geographical area, click on any news post and scroll down until you see "Discussions" preceded by the city name of your target area. This will present a list of the most recently active forum posts that contain everything from political commentary to citizens fighting with each other via text.

If you want to search for specific details by a direct URL, or if you want to bookmark your search for future use, I have found the following to be the most beneficial.

Text Search: http://www.topix.com/search/article?q=KEYWORD
User Profile: http://www.topix.com/member/profile/BILL
Location: http://www.topix.com/search/article?q=&zip=(ZIP CODE)
Google: https://www.google.com/search?q=site:topix.com+KEYWORDS

IntelTechniques Communities Search Tool (inteltechniques.com/osint/communities.html)

Similar to the previous search tools, this option attempts to simplify the various search techniques presented within this chapter. Figure 7.02 displays the current view. This tool should replicate all of the specific URLs cited within this topic. While the chances of your target appearing here are

lower than large social networks, this resource should not be ignored. In my experience, the details obtained about a target from online communities are usually much more intimate and personal than the public blasts on the larger sites.

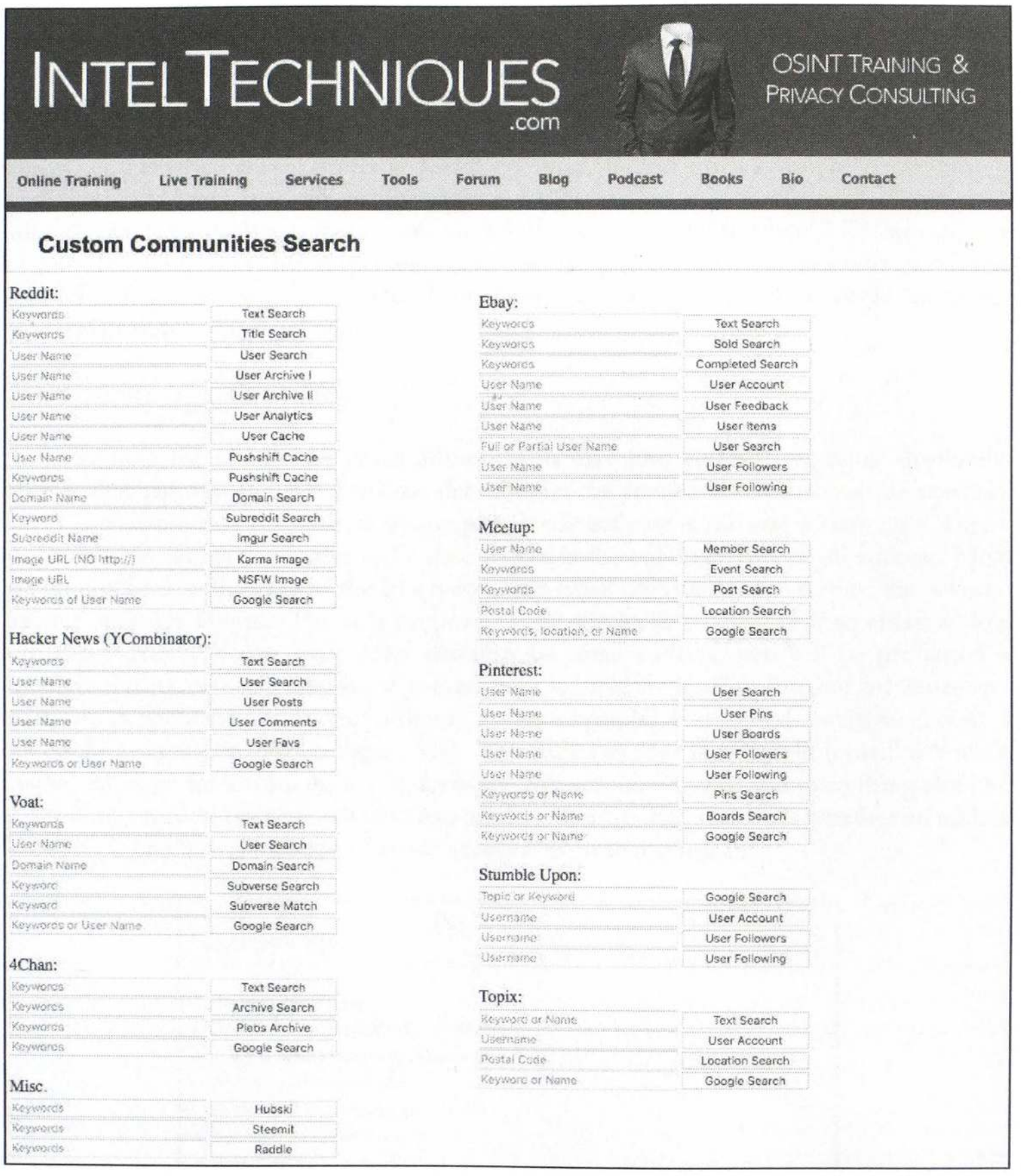


Figure 7.02: The IntelTechniques Communities Search Tool.

CHAPTER EIGHT
EMAIL ADDRESSES

Searching by a person's real name can be frustrating. If your target has a common name, it is easy to get lost in the results. Even a fairly unique name like mine produces almost 20 people's addresses, profiles, and telephone numbers. If your target is named John Smith, you have a problem. This is why I always prefer to search by email address when available. If you have your target's email address, you will achieve much better results at a faster pace. There may be thousands of John Wilsons, but there would be only one john.wilson.77089@yahoo.com. Searching this address within quotation marks on the major search engines and Facebook, as explained earlier, is my first preference. If you receive absolutely no results in these searches, you should next validate the email address.

Hunter (hunter.io/email-verifier)

When searching for a target by email address, you may find yourself receiving absolutely no results. If this happens, you need to consider whether the email address that you are searching is valid. It is possible that the address was copied incorrectly or is missing a character. There are several websites online that claim to be able to verify the validity of an email address. Most of these do not work with many of the free web-based email providers. One service that stands out from this crowd is Hunter. The sole purpose of the service is to identify if an email address is active and currently being used. After entering an email address, you will be presented with immediate results that will identify if the address is valid or invalid. Further information will identify potential issues with the address. As an example, I searched test@gmail.com, and received the results displayed in Figure 8.01. This indicates that the domain (gmail) is a webmail provider, and free for anyone to use. Otherwise, the account is valid and everything else checks out. I find this tool to be more reliable than all the others. It also identifies whether an address is a "catch-all", which may indicate a burner account from that domain.

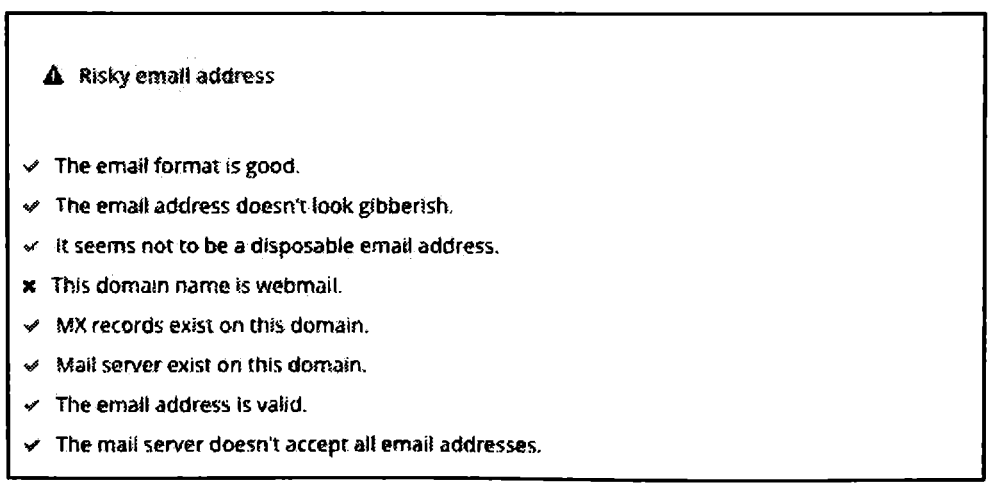


Figure 8.01: A result from Hunter.io.

If Hunter is not functioning, or is giving you inaccurate results, you may want to try **Verify Email** (verify-email.org), **TCPIP Utils** (tcpiputils.com/email-test), or an old favorite called **Email Hippo** (tools.verifyemailaddress.io). These are similar services that try to identify if an email address is valid. The results should be identical to Hunter and these services could be used to verify Hunter's response to a query. Email Hippo provides an additional feature not present in the others. As you validate your target addresses, the responses appear at the bottom as a collection. Choosing the Export option allows you to download all results as a PDF document, Word document, or Excel spreadsheet. Note that you are only allowed 20 free searches per day. In my experience, this has been sufficient for most investigations.

Find Any Email (findanyemail.net)

Find Any Email takes the technique used in the previous example to confirm valid email addresses, and attempts to discover new addresses. It asks for two pieces of information. The full name is the real name of your target. The domain field should be the internet domain of the business where your target is employed. If your target's name is John Smith, and he works at Microsoft, this service would identify the following email addresses as valid.

sjohn@microsoft.com smithj@microsoft.com jsmith@microsoft.com

Email Assumptions

You may know about one address, but not others. It can be productive to make assumptions of possible email addresses and use the verifiers to see if they exist. For example, if your target's name is Jay Stewart and he has an email address of jay112003@yahoo.com, you should conduct additional searches for the addresses of jay112003@gmail.com, jay112003@hotmail.com, and jay112003@live.com, and others. If you already know your target's user name, such as a Twitter handle, you should create a list of potential email addresses. In the next chapter, I will explain how to automate this search based on a single user name, which could also be the first portion of an email address, in order to find valid email accounts.

If I had no email address or user name for my target (Jay Stewart), but I knew that he worked at the Illinois Medical District Commission (medicaldistrict.org), I may start searching for the following email addresses.

jstewart@medicaldistrict.org
j.stewart@medicaldistrict.org
jay.stewart@medicaldistrict.org
stewartj@medicaldistrict.org

These are merely assumptions of potential addresses. Most, if not all, of them do not exist and will provide nothing for me. However, if I do identify an existing address, I now have a new piece of the puzzle to search. Creating a list of possible addresses can be time consuming. I have the following online search tool that should make this easier.

Email Permutator (inteltechniques.com/OSINT/email.html)

This tool was created after being inspired by the Email Permutator project by Rob Ousbey at distilled.net. Several similar attempts have been made to improve on the original design. One effort that showed promise was the Email Permutator+. However, modern browsers began blocking that website as malicious. I decided to create my own tool that would provide enhanced features valuable to my own investigations. This goals of this utility are to collect information about your target; create potential email addresses; and then provide hyperlinks to search for the output data. The first phase will check any custom email domain provided, as well as eleven additional popular domains. The included domains for the default "Global" setting are Hotmail, Gmail, Yahoo, Live, Hushmail, Me, Mail, Outlook, AOL, iCloud, and GMX.

The second phase of the tool is directly below the first search box. It allows you to copy the entire list of generated potential email addresses and paste them into the search box. The tool will generate direct search links for each email address through Google, Bing, and Facebook. Each link can be clicked to open a new tab to search for the chosen email address. The Google and Bing columns will each launch a quoted search of the exact email address while the Facebook links will display any profile created with the chosen address. A demonstration may explain this process better.

Figure 8.02 (left) displays the first phase of the tool including the target information. In this example, I provided a first name, last name, user name, and work domain name. The other common entries are included by default. The result was 420 potential email addresses as partially seen in Figure 8.02 (right). Notice that these results include potential addresses for multiple domains. I copied and pasted this entire list into the entry box directly below the first phase. Pressing submit generated a list of hyperlinks, as partially seen in Figure 8.03. This method can still be time consuming, but it is not as tedious as manual entry. The Russia, Germany, and Asia options will change the included domains to those most popular in each region.

This technique will work with any domain name. You can remove any domains desired in order to eliminate those results. Knowing the domain of the target's employer is a great scenario. Identifying a person's email address will often lead to their social networks, blogs, and other online accounts. I also maintain a similar tool that attempts to identify email addresses associated with a cellular telephone number. It connects the network provider's domain name used for email from the phone to the target's cellular number. It might display addresses used for social network verification messages. You can provide the results to various email search engines in hopes of connecting an account. The tool is hosted at inteltechniques.com/intel/osint/cellperm.html.

I have used this tool to generate numerous potential email addresses when only knowing my target's full name and employer. After eliminating the extended options (Gmail, Yahoo, etc), the tool creates a manageable number of links. In most scenarios, one or more of these links confirms a valid account with associated social network profiles. There are much better online tools that automate this, but none are free.



Figure 8.02: An email assumption search tool (left) and results (right).

Google Links	Bing Links	Facebook Links
jay@medicaldistrict.org	jay@medicaldistrict.org	jay@medicaldistrict.org
spice@medicaldistrict.org	spice@medicaldistrict.org	spice@medicaldistrict.org
stewart@medicaldistrict.org	stewart@medicaldistrict.org	stewart@medicaldistrict.org
jaystewart@medicaldistrict.org	jaystewart@medicaldistrict.org	jaystewart@medicaldistrict.org

Figure 8.03: Hyperlink results from an email assumption search.

Bulk Validation (leopathu.com/verify-email)

When using the previous technique to create potential email addresses, you will need a way to distinguish valid accounts from invalid. This can be done individually with the mail validation methods described earlier or in bulk with the website listed above. The bulk search will attempt to identify any valid email addresses that were created with the Permutator tool. This saves time compared to a manual search. The results here are not as reliable as with the previously mentioned Hunter, but company domains usually work well.

Email Format (email-format.com)

If the previous email assumption techniques were unproductive or overkill for your needs, you may want to consider Email Format. This website searches a provided domain name and attempts to identify the email structure of employee addresses. When searching medicaldistrict.org, it provided several confirmed email accounts under that domain and made the assumption that employee emails are formatted as first initial then last name. Our target would have an email address of jstewart@medicaldistrict.org according to the rules. I use this service to help create potential email lists from names collected from Facebook, Twitter, and other social networks. I can then verify my list with the services mentioned previously.

Online Email Databases

If you have a confirmed email address, there are numerous online tools that query multiple databases for any information associated with the account. Many of these are fronts for paid premium services that promise results and rarely deliver after payment has been received. I have found all of these to be a waste of time and money. Instead, consider the free alternatives, which are more likely to present relative results. In order from most lucrative to least, you should search your target email address through the following options.

Compromised Accounts

Email addresses are compromised regularly. Hacker groups often publicly post databases of email addresses and corresponding passwords on websites such as Pastebin. Manually searching for evidence of compromised accounts can get complicated and results are rarely complete. Several online services now aid this type of investigation. These services provide one minimal piece of information about any email address entered. It will disclose whether that address appears within any publicly known hacked email databases. While it will never disclose the owner, any email content, or passwords, it will notify you that the target's email account has been compromised at some point. It will also identify the service that was targeted during the breach. These are also good websites to check your own address. Let's start with Have I Been Pwned (HIBP).

Have I Been Pwned (haveibeenpwned.com)

This site allows entry of either a user name or email address, but I find only the email option to be reliable. The result is a list and description of any public breaches that contain the provided email address. The descriptions are very helpful, as they explain the type of service associated and any details about the number of users compromised. In a test, I searched an old email of mine that was used to create several covert accounts many years prior. The result was notification that the email address had been compromised on six websites, including Bitly, Dropbox, LinkedIn, MySpace, River City Media, and Tumblr.

Hacked-Emails (hacked-emails.com)

While Have I Been Pwned is often considered the gold standard in regard to breached account details, I now prefer Hacked-Emails for more thorough details. Using the same email address provided during the previous test, I received three additional breach notifications. These included ketquabongda.com, sendtodropbox.com, and carriermatchbox.com. When combining results from both of these services, you would now know that this target email address was likely a real account (you received results); it had been used online since 2012 (the date of the oldest breach according to HIBP); it is connected to an employed individual (LinkedIn); and it exists in spam databases as a U.S. consumer (River City Media). I believe that Have I Been Pwned and Hacked-Emails complement each other, and one should never be searched without the other. This search alone often tells me more about a target email address, even if it never identifies the owner. Note

that these services will not display sensitive results, such as whether your target had an account on the adultery website Ashley Madison. For this type of search, use Cynic (ashley.cynic.al).

Many Contacts (manycontacts.com/en/mail-check)

This premium service offers a free individual email lookup utility. It provides links to any social networks associated with the target email address. During a test search of a target's personal Gmail account, Many Contacts immediately identified the subject's LinkedIn, Twitter, Facebook, FourSquare, Instagram and Flickr accounts. Hyperlinks connect you directly to each profile page. This is one of my first searches when I know I possess a target's primary email account.

Pipl (pipl.com)

Enter any email address and you will be presented a dossier style of report of all available information. Much of this data has already been received from other sources. However, the user name portion often includes unseen details, which may include associated subjects or non-relatives that do not appear in other search results. If you want a more detailed report, consider the Application Programming Interface (API) version of Pipl that is explained in Chapter Twenty.

Gravatar (gravatar.com)

This service is responsible for many of the small image icons that you see next to a contact in your email client. You may notice that some incoming emails display a square image of the sender. This is configured by the sender, and this image is associated with any email address connected. You do not need to wait for an incoming message in order to see this image. While the Gravatar home page does not offer an email address search option, we can conduct a query directly from the following URL, replacing the email address with your target information. This image can later be searched with a reverse image query as explained in Chapter Fourteen.

https://en.gravatar.com/site/check/test@gmail.com

That's Them (thatsthem.com)

The majority of the email addresses and user names I have searched through this service returned no results. However, on occasion I received detailed results such as full name, address, phone number, and vehicle information. Although this is rare, I believe That's Them should be on your list of email and user name search resources.

Domain Connections

Every domain name registration includes an email address associated with the site. While many people will use a privacy service to hide personal details, this was not always the case. Fortunately, many free services have been collecting domain registration details and offer queries of current

and archived domain registration data. They will identify domain names that were registered with the target email address entered. This is beneficial when you have a tech savvy target that may have registered websites of which you are not aware. This also works on domains that no longer exist. As a test, I provided an email address of brad@notla.com. The following results identify the details found by each service.

Whoxy (whoxy.com/reverse-whois)
Full Name
Home Address
Telephone Number
11 Domain Names
Registrars and Hosts

Domain Big Data (domainbigdata.com)
Full Name
Home Address
Telephone Number
8 Domain Names

DNS Trails (dnstrails.com)
7 Domain Names

Whoismind (whoismind.com)
Gravatar image
3 Domain Names

Analyze ID (analyzeid.com)
Amazon Affiliate ID
AdSense Affiliate ID
5 Domain Names

The use of affiliate IDs obtained by services such as Analyze ID will be explained during the domains instruction presented in Chapter Sixteen. Additionally, some of the most impressive email address search options I have found are only available by using a service's application programming interface (API), as explained in Chapter Twenty.

Imitation

The ***assumptions*** of email addresses can be valuable in discovering potential valid accounts, as mentioned earlier. ***Imitation*** of any target email addresses can reveal more details, and confirm association with online activities. Consider the following example. Your target email address is bill@microsoft.com, and you want to know if he is a Mac or Windows user. You could first navigate to appleid.apple.com/account and attempt to make an Apple account with that address.

If allowed to proceed past the first screen, then that user does not already have an account there associated with the target address. If you are informed that the email address is already in use, then you know that your target is a Mac user, and that specific address controls the account. You could then navigate to signup.live.com and attempt to create an account with the target address. If denied, then you know that your target is already a Windows user, and the supplied address controls the account.

This method can be replicated across practically all websites and services. I have used this technique to confirm that target email addresses are associated with services from Yahoo, Gmail, Facebook, Twitter, and many others. I have also identified real-world services in use by my target by attempting to create accounts with local cable, power, and water companies, supplying the target email account and being notified that the address was "already in use". Knowing the local cable provider of a suspect can seriously limit the geographical area where he or she could be residing. Be sure not to fully execute any account creations with your target email address, as he or she will receive a notification. This method should only be used on an initial account creation screen, and never submitted.

Email Provider

If your target's email address ends in gmail.com or yahoo.com, the identity of the email provider is quite obvious. However, business addresses and those with custom domain names do not notify you of the service that hosts the email. A domain's email provider is the company listed in the domain's MX record. The email provider may be the same as the domain's hosting company, but could also be a separate company. You may need to know the email provider in order to issue a court order for content or subscriber data. You may simply want to document this information within your investigation for potential future use. Regardless of your needs, the following will obtain the email provider from almost any address.

Navigate to **MX Toolbox** (mxtoolbox.com) and enter the domain of the email address, such as phonelosers.org. The result should include a hostname and IP address. These identify the email provider for the target domain. In this example, the host is mx1.sub4.homie.mail.dreamhost.com. This tells me that Dreamhost is likely the email provider. You could have also replicated this process in one click using the email search tool at **Records Finder** (recordsfinder.com/email).

These techniques help me identify the email providers or hosts behind business email accounts. If I am conducting a consensual penetration test, this information may lead me toward a social engineering attack against the host. If I am trying to connect an individual to a shell company, this may associate the same small provider with each target. I believe every thorough OSINT report should include a brief mention about the domain email provider. This should be checked as the investigation continues. Changing providers could be a sign of paranoia or intent to conceal evidence. Law enforcement can use this information in order to secure a proper search warrant.

IntelTechniques Email Address Search Tool (inteltechniques.com/osint/email.search.html)

Similar to the previous IntelTechniques search tools, I have created a custom email address search page. This website is basically a shortcut to many of the email search tools that were presented here and in previous chapters. The Populate All and Submit All buttons will automate the process of searching an email address through Hunter, Google, Bing, Facebook, Hacked-Emails, Have I Been Pwned, Pipl, That's Them, Reverse Mails, Google Newsgroups, Google Blogs, FTP Servers, Domain Big Data, WhoisMind, DNS Trails, Analyze ID, and Gravatar. Figure 8.04 displays the current state of the page.

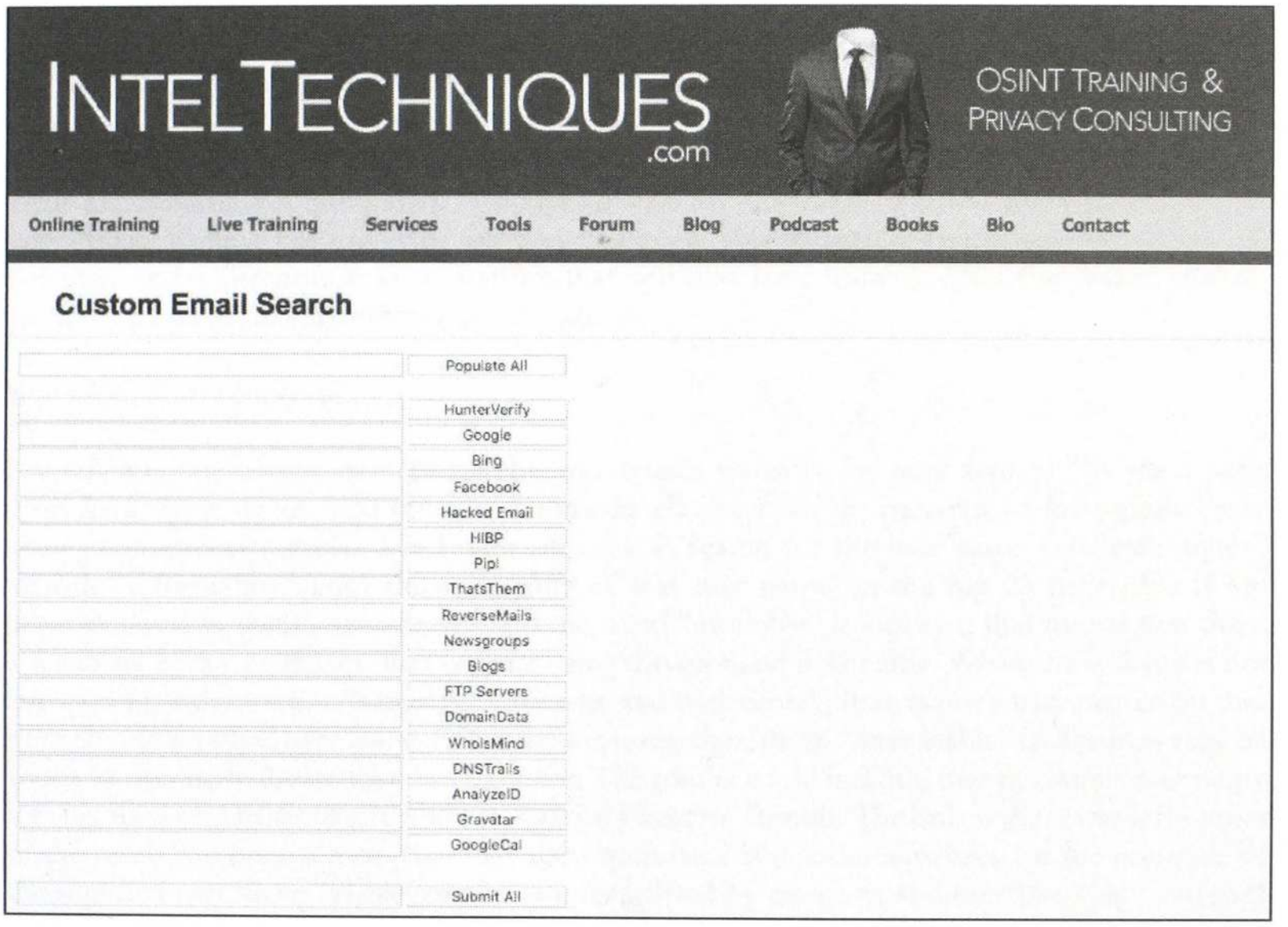


Figure 8.04: The IntelTechniques Custom Email Search Tool.

CHAPTER NINE
USER NAMES

Once you have identified a user name for an online service, this information may lead to much more data. Active internet users often use the same user name across many sites. For example, the user "amandag62002" on MySpace may be the same "amandag62002" on Twitter and an unknown number of other sites. When you identify an email address, you may now have the user name of the target. If a subject uses mpulido007@gmail.com as an email address, there is a good chance that he or she may use mpulido007 as a screen name on a number of sites. If the target has been an internet user for several years, this Gmail account was probably not the first email address used by the target. Searches for mpulido007@yahoo.com, mpulido007@hotmail.com, and mpulido007@aol.com may discover new information. Manual searching of this new user name information is a good start. Keeping up with the hundreds of social websites available is impossible. Visiting the following services will allow you to search user names across several websites, and will report links to profiles that you may have missed. After the details of each service, I provide a comparison chart of features.

KnowEm (knowem.com)

KnowEm is one of the most comprehensive search websites for user names. The main page provides a single search field which will immediately check for the presence of the supplied user name on the most popular social network sites. A search for the user name "inteltechniques" provides information about the availability of that user name on the top 25 networks. If the network name is slightly transparent and the word "available" is stricken, that means that there is a subject with a profile on that website using the supplied user name. When the website is not transparent and the word "available" is orange and underlined, there is not a user profile on that site with the supplied user name. For an online researcher, these "unavailable" indications suggest a visit to the site to locate that user's profile. The results could indicate that the target user name is being used on Delicious and Twitter, but not Flickr or Tumblr. The link in the lower left corner of the result will open a new page that will search over 500 social networks for the presence of the supplied user name. These searches are completed by category, and the "blogging" category is searched automatically. Scrolling down this page will present 14 additional categories with a button next to each category title stating "check this category". This search can take some time. In a scenario involving a unique user name, the search is well worth the time.

Name Chk (namechk.com)

Name Chk provides a similar service. It does not search as many sites as KnowEm; however, it provides a feature that is a great convenience for the researcher. Entering a unique user name in the search field at the top of the page will immediately identify the presence of that name within 118 popular social networks. A green background indicates that the user name is not in use on

that site while a dark background indicates that a user account exists on the site. The advantage with this site is that clicking on Download Results presents a CSV spreadsheet with raw data.

Check User Names (checkusernames.com)

This site combines the search engine of KnowEm and the features of Name Chk. It searches approximately one third of the sites on KnowEm, but it links directly to the target's profile when one is identified. Because this service relies on KnowEm, and not its own internal structure, it often experiences outages.

Name Checkr (namecheckr.com)

This service appeared in late 2014 and conducts the same type of search as the previous competitors. The only slight advantage here is that the search is conducted faster than other sites. Additionally, you have a live hyperlink that will navigate to any identified accounts of the target.

User Search (usersearch.org)

This service stands out a bit from the others in that it only provides actual profile results. It searches the supplied user name for a presence on 45 of the most popular websites (basic option) or 115 total websites (advanced option) and returns a list of identified profiles matching the target. While this service is the slowest of all options, this could be an indication of account verification for more accurate results. I have also found their email address search occasionally valuable.

NameVine (namevine.com)

This user name search service provides a unique feature missing in the rest. It allows you to begin typing any partial user name and it will immediately identify registered accounts within the top ten social networks. This could be beneficial when you are not sure of the exact name that your target is using. If your target has a Twitter user name of "Bazzell", the previous services will easily identify additional accounts that also possess this name. If you think that your target may be adding a number at the end of the user name, it could take some time to search all of the possibilities. With NameVine, you can quickly change the number at the end of the user name and get immediate results. It will search Twitter, Facebook, Pinterest, YouTube, Instagram, Tumblr, Wordpress, Blogger, and Github. It will also check for any ".com" domains that match. The benefit of this service is the speed of multiple searches.

Pipl (pipl.com)

Pipl has been discussed as a great site for searching a person's real name and email address. This site performs equally as well at locating people by a user name. Inserting a user name in the same field that a person search would be conducted will display results of subjects using this user name on social networks. It will also attempt to determine vital information about the user including

age, location, employer, and interests. Finally, it will display small images that are associated with the user's social network accounts. An important part of searching by user name is the attempted searches of unknown user names. In all of the examples above, the user name "inteltechniques" was used. If you know that your target is using this name, you may want to take a quick look at the user names of "inteltechniques2", "inteltechniques3", "inteltechniques4", etc. While these profiles may not belong to your target, you could discover new profiles that would otherwise have been missed.

Peek You (peekyou.com)

This service has recently introduced new search options and better accuracy. The standard landing page encourages a search of a person's first and last name. A filter by country option exists which may eliminate unwanted results. This often identifies a target's Twitter page, Facebook profile, and related accounts. This basic data is only the beginning of the service's offerings. The "User name" search option performs a query similar to KnowEm that will identify social networks that possess a user account with the specified user name. On occasion, this has located new internet profiles that other services missed. The "Work" search option attempts to locate people by their employer.

Lullar (com.lullar.com)

Lullar will search by email address, user name, or real name. The search excels with an email address or screen name, but I do not recommend the real name option for reliable results. Lullar takes a different approach with the search results. When conducting the search, the results page appears almost immediately. This is because Lullar is not actually conducting any real analysis of user profiles. The results displayed are only the links that would open the page of the target's profile based on the address, or URL, of that profile. For example, if I search for the user name JohnDoe911, Lullar does not check any sites to see if this user has a profile. Instead, it generates the appropriate links that would function for that user name. In the case of Twitter, it creates a link to twitter.com/#!/ search/JohnDoe911. This link will be presented whether there is a profile at this address or not. On the down side, you will often be presented with links that do not function. On a positive note, you may get links that do function and are so new that other engines have not indexed them yet.

I often use this service in two scenarios. When I have a partial or questionable user name, I will search it through Lullar and see what the profiles look like. This can tell me right away if I am researching the wrong target name. The other scenario is when I encounter an email address or user name that may have alternatives or aliases. If my target has a user name of TheJohnDoe2, I am curious if the user also possesses TheJohnDoe3, TheJohnDoe4, or maybe TheJohnDoe. Many times, users will need secondary user names and will choose names very similar to their primary names. Lullar will show me what content appears on these profiles without regard to the real name of the user of the accounts. This may lead to unwanted profiles, but has been very successful at locating previously unknown profiles of a target.

User Sherlock (usersherlock.com/usersearch)

While I NEVER recommend using this site's email search feature (it discloses search attempts and your approximate location to the email target), the user name option is relatively safe. The only benefit here is the ability to see extended information about each positive result. Real names, locations, and direct links are presented within the minimal results offered. Most of this will be redundant details from the previous queries.

Ultimately, you will need to determine which of these sites works best for your preferences. I recommend using all of them as necessary. I believe that KnowEm is the best start when you have an absolute user name, while NameVine helps when you are still trying to identify a user name. The following chart identifies the number of networks searched, whether the results connect directly to the target page, and any additional features.

Service	# of Sites	Active Links	Features
KnowEm	596	NO	
NameCheck	106	NO	CSV Download
CheckUserNames	160	NO	
NameCheckr	44	YES	Checks Domain
User Search	115	YES	
NameVine	10	YES	Live Results
Pipl	Varies	YES	Positive Results
Lullar	25	YES	
User Sherlock	20	YES	Extended Details

Snapchat (findmysnap.com)

An exploit with the Snapchat user database was explained in Chapter Four when discussing Facebook cellular telephone searching. Find My Snap maintains an online search tool that queries any user name within this data set and identifies any leaked information. For many user names, this tool will display the first eight numbers of a ten-digit telephone number. While not complete, it can identify the area code and prefix of a user name, which may identify an approximate location such as city and state.

Skype User Name (web.skype.com)

Identifying a Skype user name can be an important lead. It could direct you toward additional searches of the newly found data. Unfortunately, user names are not obviously available when researching a real name on Skype. However, a quick method will reveal any Skype user name when searching by name or email address. While logged into a Skype account within the website or application, navigate to the search area. This section allows you to enter a real name or email address, and then conducts a search of the Skype user directory. Any results will appear immediately below. Clicking on these results displays the user's basic profile details including a

photo. If the user chose not to include a photo, a silhouette graphic appears. Right-click on either image format and choose to "Open image in a new tab" (Chrome) or "View image" (Firefox). The new tab will contain the image that was already available. However, the address in the URL will reveal the Skype user name of the target. The following URL was created from the profile image of a Skype user with an email address of lorangb@gmail.com. It identifies the target's Skype user name as bart.lorang. The user name will always be between two forward slashes (/) after the word "users" within the URL.

https://api.skype.com/users/bart.lorang/profile/avatar?cacheHeaders=1

If your target is a Skype user, this technique might translate an email address into a user name that could be present across multiple networks.

IntelTechniques User Name Search Tool (inteltechniques.com/OSINT/username.html)

Similar to the custom email search tool, this page assists with an automated search of some of the techniques mentioned previously. This page, seen in Figure 9.01, allows you to enter a single user name and populate all of the search options. You can then execute manual queries or use the last option to attempt all search possibilities. As a reminder, you must allow pop-ups for the domain of IntelTechniques.com if you want to use the "Submit All" option. This tool currently searches your target name through KnowEm, NameVine, Check User names, Pipl, PeekYou, User Sherlock, UserSearch, Twitter, Facebook, YouTube, Tumblr, Instagram, Google+, and Email Search. The last option populates a search on Google of the target user name with the addition of the most popular email address domains, including quotation marks and the OR operator, as discussed in Chapter Three. If your target user name is IntelTechniques, it conducts the following Google search.

"IntelTechniques@gmail.com"OR"IntelTechniques@yahoo.com"OR"IntelTechniques@hotmail.com"OR"IntelTechniques@live.com"OR"IntelTechniques@outlook.com"OR"IntelTechniques@comcast.net"OR"IntelTechniques@verizon.net"OR"IntelTechniques@charter.net"OR"IntelTechniques@aol.com"

If desired, you could copy this query and paste it into Bing and Yandex, but I have found the results to be very unreliable. This search page is always my first stop when I have a target user name. It will usually generate the most results within the quickest amount of time.

Compromised Account Assumptions

In the previous chapter, I explained how I use Have I Been Pwned (haveibeenpwned.com) and Hacked Emails (hacked-emails.com) in order to identify online accounts associated with a target email address. These services do not work properly with user names, and entering a user name does not check it against related email addresses. Similar to the previous instruction, we can make assumptions in order to identify target accounts. Assume that your suspect is IntelTechniques on

Twitter. A manual search of IntelTechniques@gmail.com, IntelTechniques@hotmail.com, and IntelTechniques@yahoo.com at Hacked Emails might reveal an active account that appears within a compromised database. Unfortunately, this manual search is time consuming and takes a lot of redundant effort. Therefore, consider using my previously mentioned custom search tool.

Figure 9.01 displays four search fields near the bottom of the page. The first two allow entry of a user name, and then populate several tabs on your browser, each searching for that user name at a popular email domain. The last two search fields replicate this search, but display the results within the search page for immediate viewing. This is all conducted through the application programming interfaces of Have I Been Pwned and Hacked Emails, and the exact technique will be explained in Chapter Twenty. Figure 9.01 displays results after conducting a search for the user name "mikeb". The results in this example confirm that the email addresses mikeb@gmail.com, mikeb@yahoo.com, mikeb@hotmail.com, and mikeb@live.com all appear within compromised account databases. This indicates that these addresses were active at some point in time, and were associated with the specified online services which experienced data breaches. This is a quick way to determine if a unique user name likely exists within a possible target email address. If I had searched IntelTechniques4 and received a positive hit on the email address IntelTechniques4@yahoo.com, I would want to research that email address further using the techniques in the previous chapter. I would also want to identify any account details present on the service that experienced the breach.

Similar to how compromised database searches are the most powerful email search option that I use, querying user names in this manner can be equally important. This search tool eliminates any laborious process and removes any excuses not to conduct this type of search every time. The "Populate All" button at the top of the page will repeat your target data into all of these fields for easy execution.

Real World Application: In 2017, I was investigating a subject that had stolen a large amount of money and disappeared. A private investigator determined that the identity provided during the transaction was an alias, and no real name was known. The suspect provided an email account similar to BradMartin765@gmail.com. Using the search techniques mentioned in the previous chapters, I discovered that BradMartin765@gmail.com was connected to a Facebook page created in the same alias name. However, this profile possessed a user name similar to crazycheetah70. A search of crazycheetah70 on my search tools page revealed that crazycheetah70@yahoo.com was a valid email address, and had been compromised during the LastFM data breach several years prior. Navigating directly to the profile page of last.fm/user/crazycheetah70 indicated that the account had been deleted. However, a search for that URL on the Wayback Machine, as explained in Chapter Three, revealed the archive of this profile, as seen before deletion. Enough true data about the target was available in this profile to determine the real identity. While these connections of accounts and recycling of user names was very careless, most suspects make sloppy mistakes eventually.

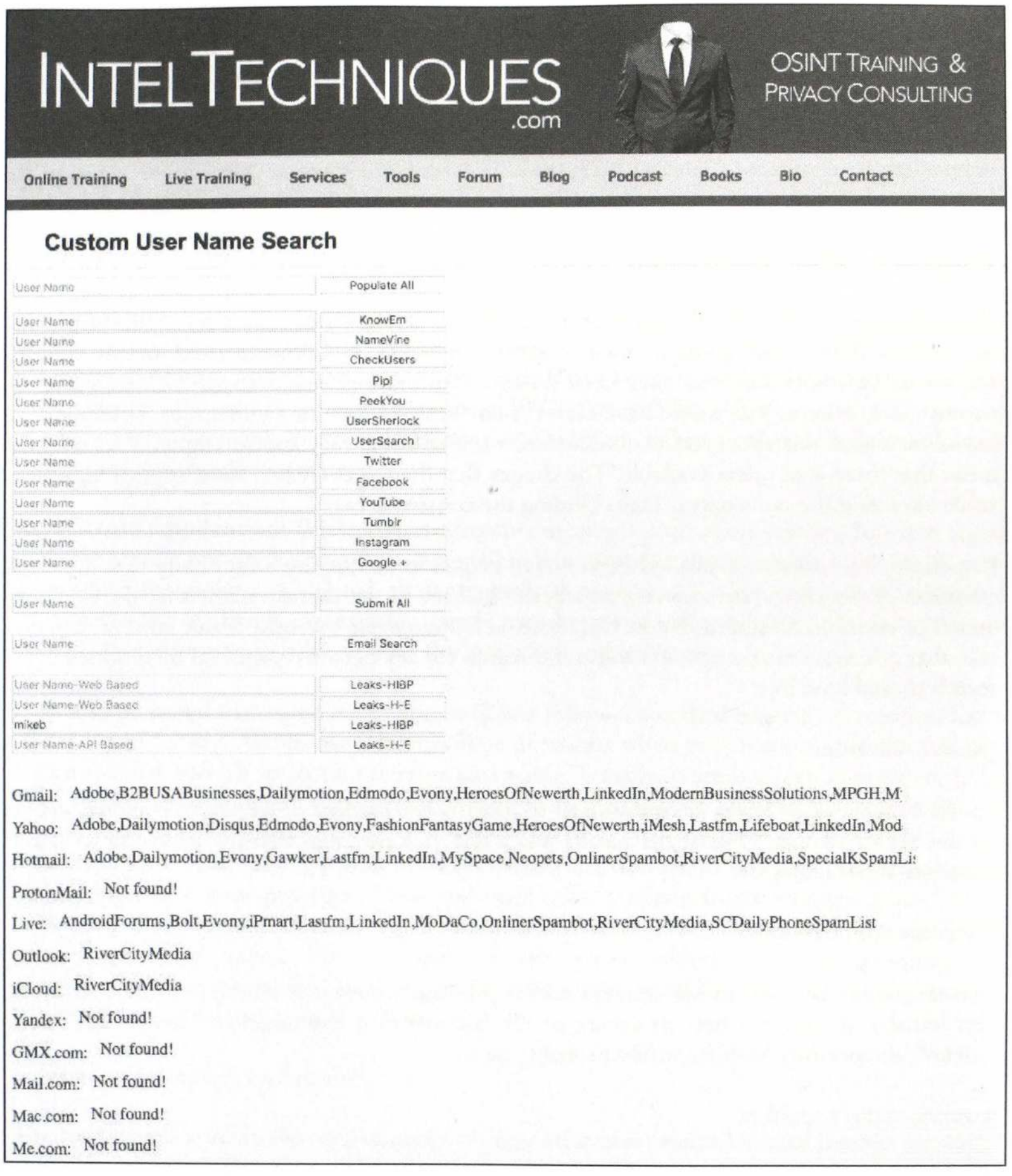


Figure 9.01: The IntelTechniques Custom User Names Search Tool.

University Homepages

These automated searches for user names can be very productive. However, they will not locate accounts within all online communities. One large untapped resource is the massive presence of university personal web pages and user names. Almost every university issues each student a university email address. These usually follow a standard naming convention such as the following.

lastname.firstname@university.edu

If you can identify the convention that the school uses and know the full name of the target, you can determine the email address of the student. This address can be searched for any websites and social networks that may have been missed. Furthermore, the first part of that address is usually the user name that would have been issued to the student for a homepage. The target may have never used this email address online and a search result may appear empty. That does not mean that there is not data available. The chance that the target created some type of homepage while attending the university is high. Finding the content is easy.

Hopefully, the searches explained earlier have helped in identifying a university that the target attended. A search for the university's website will reveal the domain name that the university uses. For example, Southern Illinois University at Edwardsville's website is siue.edu. We can now take that information and conduct a specific search for any personal pages on their domain. The search should look like:

site:siue.edu laura

I picked the name of Laura at random just to identify any student or employee personal website on the SIUE domain. One of the results was a link to a personal website belonging to "Laura Swanson". The link was:

www.siue.edu/~lswanso/

This indicates that the naming convention for personal websites is a tilde (~) followed by the first initial and then the first six letters of the last name. If the target of interest was "Scott Golike", the personal website would probably be at:

www.siue.edu/~sgolike/

We can also assume the possibility of his school issued email account to be sgolike@siue.edu. A few searches using previously discussed techniques should confirm if this address belongs to the target. A search using the email to Facebook profile technique may identify an abandoned profile.

We can now navigate to this personal school page and see if there is any content. If there is, we can collect the data and conduct an analysis for intelligence and further research leads. If there is no page at this address, it does not mean that there has never been data there. This only indicates that there is no current content on this website. When students graduate, universities will usually remove all of the personal content from the servers. As discussed previously, this is never an excuse to stop looking. You can now take the URL of a target and conduct a search on The Wayback Machine (wayback.archive.org).

As an example, I can navigate to the first personal link for "Laura Swanson". Figure 9.02 displays a portion of the live page at www.siue.edu/~lswanso/. If this page did not exist and the site contained no content, you could check on The Wayback Machine. Figure 9.03 shows the search results for this personal page and identifies numerous archives dating back to 1997 for this site. Checking all of these options presents the many different versions of the site including one from 2005 (Figure 9.04) and the first capture from 1997 (Figure 9.05). This presents new data that would not have been uncovered with conventional searches. When a personal website is located, earlier versions should be archived.

Real World Application: While assisting another agency, information had developed in regard to a suspect in a priority investigation. After all online search attempts revealed nothing of value in locating the subject, a deleted student personal page was located using this method. It contained a list of friends, roommates, family members, and interests that were not previously known. This information helped locate the individual within hours.

It should be noted that some institutions will not follow a standard naming convention for all students and faculty. Additionally, there will be occasions when two or more students will have a name similar enough to create the same user name. Usually, there is a plan in place to thwart these duplications. Sometimes it is as simple as adding the number "2" after the user name.

Universities are not the only places that create personal web pages based on a member name. Several internet service providers allow each subscriber a personal space online as part of the provider's main website. Comcast provides 25MB of storage in a folder with the title of the subscriber's user name. For example, if the email address of the customer was crazycheetah70@comcast.net, the user name for the service would be crazycheetah70. The URL to view the personal web page would be:

home.comcast.net/crazycheetah70

The following is a sample list of personal web page addresses from additional internet providers, using "crazycheetah70" as a user name example. You should also search for internet providers in the target's area and attempt to find deleted pages on The Wayback Machine, Google Cache, Bing Cache, Yandex Cache, and the other services discussed in Chapter Three.

360.yahoo.com/crazycheetah70
crazycheetah70.webs.com
crazycheetah70.weebly.com
webpages.charter.net/crazycheetah70
sites.google.com/ crazycheetah70
about.me/ crazycheetah70
angelfire.com/ crazycheetah70
geocities.com/ crazycheetah70
reocities.com/crazycheetah70
crazycheetah70.tripod.com
home.earthlink.net/~crazycheetah70
home.comcast.net/~crazycheetah70

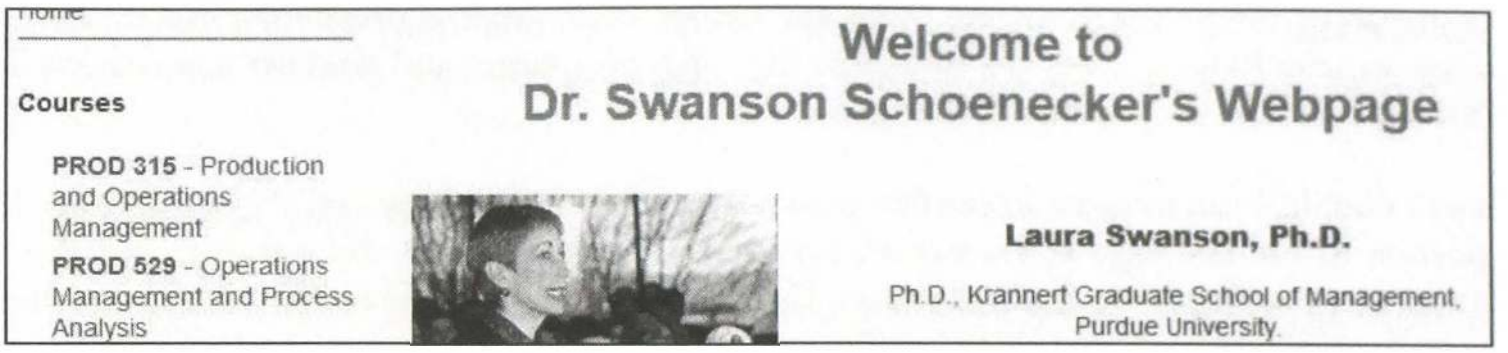


Figure 9.02: A current personal page on a university website.

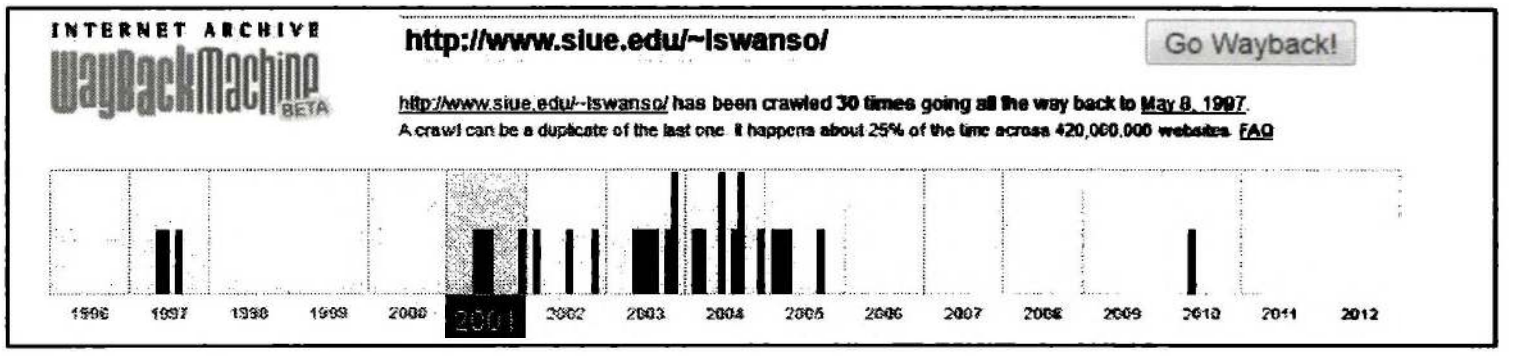


Figure 9.03: A Wayback Machine timeline of available versions of a website.

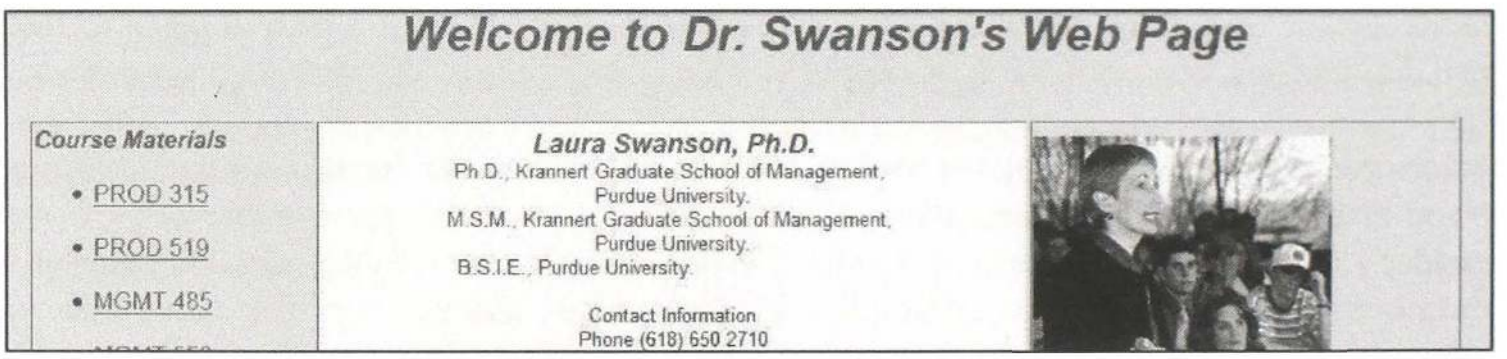


Figure 9.04: A previous version of the website.

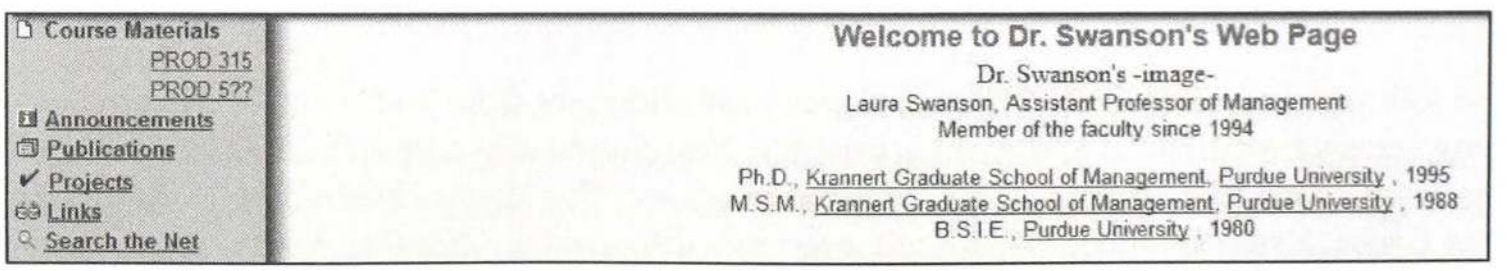


Figure 9.05: A previous version of the website.

Chapter Ten
People Search Engines

Just as Google and Bing specialize in searching content on the internet, people search engines specialize only in finding content about a particular person. Many of these sites utilize search engines such as Google and Bing to help compile the data, and then present a summary style interface that is easy to consume. The sites listed here each have their own strengths and weaknesses. Standard searches are free on all of them; however, each site generates revenue in some form. Usually, this is by displaying advertisements that often appear to be a report function within the site. I do not recommend purchasing any of the premium paid services until all free options have been exhausted. These details are often focused on targets in the United States, but many services are starting to reach past North America. Searching a target's real name often leads to the discovery of home addresses, telephone numbers, email accounts, and user names. After an explanation of each service's offerings, I will present a comparison of the data available. This will help display the specific types of free information available at each site.

Pipl (pipl.com)

This site claims to be "the most comprehensive people search on the web". Entering a first and last name along with a city and state associated with the target will generate a new page full of information. If you are not sure of the location of your target, just enter a first and last name. The first group of data you will see identifies subjects matching your search along with location information for each. If any appear to reside in a general area that is near your target's last known location, you should investigate the link. The next group will have icons of images from various social networks. Each of these will link to the original image location, which will be a part of a user's social network profile.

The results will present links to social network profiles that belong to users with the name that you searched. Common networks include Twitter, Facebook, Meetup, MySpace, and YouTube. The next section will include web page hits on your target name. These are usually fairly accurate to the target name supplied. These tend to place a stronger emphasis on personal websites and blogs than general pages that a standard search engine might provide. Ultimately, with Pipl, you want to identify your actual target through a Pipl profile. To do this, you need to filter through the results you receive. As an example, I conducted a search for a name without a location. The results provided the location information for the target. This page links to a Pipl profile with a lot of information about the target.

The left column of this page provides suggested filters for location and age. This will help filter the results in the case of a common name. Once you locate the Pipl profile for your target, it will display all available social network associations that Pipl has on the target. This will include more

photo icons that link to the host profile. Many of the links on this page will be "Sponsored" links that will direct you to an advertiser's website. Fortunately, Pipl marks all of these links with the word "Sponsored". I stay away from these traps. Often the link will provide enough visual information that clicking for further data is unnecessary. Sometimes an entire telephone number, age, or family member's name is visible without visiting the sponsored site. The classmates section, a sponsored link, will always identify the school attended, years attended, and location of the school in the text of the link without actually clicking on the link. The rest of the page will vary depending on how much information is available about your target. Pipl will never present all data that is out there about your target. It is simply a worthwhile stop to begin identifying related content. The information obtained here can be valuable for future searches. Pipl also allows searching through their API, which will be explained later in this book.

That's Them (thatsthem.com)

In late 2014, a new website quietly entered the crowded scene of people search services. On the surface, it was just another service that aggregated publicly available information. Consequently, a closer examination revealed That's Them to contain information that is not available anywhere else for free. This service has many options, and most will be discussed in this book. For the purposes of this chapter, I will focus on the "Name and Address" search option in the top menu of the website. Entering a full name with city and state is preferred, but not required. Results often display the person's age range, cell phone number, landline number, full address, religion, financial details, home IP address and any associated email addresses. I searched my own name to test the accuracy of the results. My profile correctly identified similar information as well as the exact VIN number of a previous vehicle. This type of data is impressive without any fees. I have found their details of religion and financial information to be unreliable. Note that the options to purchase additional information are advertisements from third-party companies, and should be avoided.

Spokeo (spokeo.com)

Spokeo is probably the most well-known of all of the people search engines. There are two very distinct versions of this service, free and premium. The premium service will provide a large amount of accurate data, but at a cost. The free version provides an interface that is easy to navigate. The results from a target name search will be presented after choosing a state and city. Only the states and cities where the target name has a presence will be shown. Choosing this target will display a profile with various information. Within this data will be several attempts to encourage you to purchase a premium account. Basically, anything that you do not see within this profile will cost you money. Any links from the profile will present a membership plan with pricing. The profile will often display full name, gender, age, and previous cities and states of residency. However, it no longer presents the actual current address. Well, kind of...

As an example, I searched for people with my name and located an exact match. The profile page only identified a city in West Virginia as a home address. However, after right-clicking on the

page and choosing the option to "View Page Source", I could see the following details (modified for privacy) within the text. While Spokeo will not spoon-feed residential addresses to you without paying, the data that you need is within the source code of each profile.

meta property="og:description" content="Spokeo profile page for Michael Bazzell in Green Spring. Michael Bazzell, age 38, lives at 6746 Chester Valley Rd.

Advanced Background Checks (advancedbackgroundchecks.com)

This is another service that includes advertisements for premium options. Surprisingly, the majority of their data archive is free without any payment. The main search results appear redacted and entire addresses and telephone numbers are masked. With many services, clicking these details prompt the user for payment. Instead, this service opens a new page revealing the entire record. This often includes home address, home landline telephone number, age, and relatives. Clicking the "See full info" button reveals previous addresses, additional telephone numbers, and aliases. Overall, this service is extremely useful for U.S. targets.

Yasni (yasni.com)

On the surface, Yasni appears to be another standard people search engine. Much of the content received will be duplicate data, but there are a few areas where Yasni works differently. The home page will give three search options. For most OSINT purposes, the last option is the desired search. It will accept a real name or user name and forward you to a results page. Real name search will present a large number of links associated with your target's name. As with other engines, many of these results will be about a person other than your target. The first box on the results page will include a "lives/works in" option that will display the cities of the users identified with the search. Clicking on a location that looks appropriate for your target will load a new results page that will provide all search results about your specific target. These links could all be found using search engines and operators, but this will take the hassle out of that technique. Though to obtain complete results on a target, you should still visit a standard search engine. This Yasni page will identify news articles, websites, and social networks related to your target. By default, the search is conducted internationally. Yasni is a German site and searching outside of the United States is one of the strengths of the service. The search bar includes an option to filter the results by specific countries, but the United States is not listed as an option. If you have a target that lives in another country, Yasni is a great tool.

Intelius (intelius.com)

Intelius is a premium service that provides reports about people for a fee. Most of the information is from public sources, but some of it appears to come from private databases. Searching for any information on the main website will always link you to a menu of pricing options. The information will never be displayed for free. However, the page that lists the report options does possess some interesting information. This free preview identifies an exact age, possible aliases,

cities lived in, previous employers, universities attended, and relatives. If the subject is married, this will usually identify the spouse. In most situations, it will identify the maiden name of the person's wife. Anything that you do not see on this main screen, you must pay a fee. I never recommend purchasing any of this data. Users are usually disappointed with the results.

Radaris (radaris.com)

This service has many similarities to Intelius. However, the data set is unique. The business model is to entice you into purchasing an entire profile. I only use the service for the limited free content available in the preview. After searching a name, select the most appropriate target and choose "Full Profile" in the lower right of the result. This will open the full view of any free information. This will often include the target's middle name, age, current address, previous address, landline telephone number, and links to social networks. The Background Check options will forward you to a third-party premium access website that I do not recommend.

Zaba Search (zabasearch.com)

This site appears to have several search options at first glance. Unfortunately, all but one will forward to an Intelius site, which will require a fee. Though there is one very specific free option on this page. Providing any real name and state of residence will provide a results page with full name, date of birth, address and phone number. In my experience, this often includes unlisted telephone numbers and addresses. Clicking on practically anything else on this results page will take you to a sponsored link. When I use this resource, I only rely on the information obtained on the first result page.

Melissa Data (melissadata.com)

This service provides data on many U.S. and Canadian households. The results within each profile will appear impressive at first glance, but most of the data is generic to the geographical area near the target's home. The most value in reference to a target will be in the general results page. This list should identify full name, home address, and age of your target.

Fast People Search (fastpeoplesearch.com)

This service appeared in 2017 and possesses a staggering amount of information on U.S. individuals. A typical profile includes the full name, age, and home address of your target. From there, multiple email addresses and telephone numbers populate the majority of the results page. This historic data has been extremely valuable in my investigations. Previous addresses include dates of residency, similar to what you would see in a premium service. Relatives and associates appear reliable in my experience. If your target is on this dataset, you may not need any other search options listed here.

Nuwber (nuwber.com)

Another newcomer in 2017 was Nuwber. I first learned about this service from various members of the privacy forum at my website. They were discussing the importance of removing their own personal details from this site through various opt-out procedures available at the time. I found my own information to be quite accurate. Therefore, this makes for a great OSINT resource. The default landing page allows search of a first and last name. The results are presented by location, and each profile often includes full name, age range, home address, telephone number, and neighbors.

Find People Search (findpeoplesearch.com)

Similar to the others, this website allows search of a first and last name, and presents profiles with further details. The strength of this service is the presence of historic home addresses, telephone numbers, email addresses, and relatives. While much of this data is no longer accurate, the past details are often very helpful with background checks.

Family Tree Now (familytreenow.com)

In 2016, this website emerged and launched an uproar online. This site is targeted toward those that want to conduct family history research, and its specialty is connecting a person to his or her relatives. The results do not display home addresses, but simply the age of the target and a large list of family members. After gaining a lot of online popularity, many people started complaining about the availability of this sensitive information. While this type of service is nothing new, people were outraged at this violation of their privacy. Since Family Tree Now sources all of their data from public databases, they defended their product, which is still available today. This attention may have been the inspiration for this company to take things to another level with True People Search.

True People Search (truepeoplesearch.com)

Similar to the previous option, True People Search allows queries of real names. However, the results are quite different. While the relative and associate data might be the same as Family Tree Now, these results also include current home address, current telephone numbers, previous home addresses, previous telephone numbers, email addresses, and dates connected to each. This is another impressive data set. Surprisingly, there was little attention given to this new product. One advantage that researchers might have here is that those who requested removal of their personal details from Family Tree Now likely never thought to repeat the process on this site.

People Search Now (peoplesearchnow.com)

This database appears to have the same parent company as True People Search, and possess the same data. However, this search should be included in the event that a target has removed details

from other related websites. Due to increased awareness of exposed personal information, I am seeing many people request removal of personal online data. A new site called **John Doe** (johndoe.com) also appears to use the same data, and displays opt-outs from the other sites.

Cubib (cubib.com)

This was another service that surprised many privacy-conscious people. A typical entry contains a target's full name, current home address, current home telephone number, email addresses, previous home addresses, additional telephone numbers, possible relatives, and age. I recently located a target's email address from this service, which was not present on any other site. Using the techniques mentioned in previous chapters, I was able to create a full dossier about him.

WP Numbers (wpnumbers.net)

This service sources much of their details from public domain registrations, similar to the services discussed in Chapter Eight. The difference here is the ability to search by real name (instead of email address). A search of my own name revealed a specific address and telephone number that I used to register several websites many years prior. While this data may not be the most accurate details about your target, it could present an option not seen within any of the other resources.

Peek You (peekyou.com)

This site offers real name, user name, and telephone number search options. The real name search can be extremely useful. The typical results include home address, age, telephone number, and associates. The additional layer with this service is the presence of a social network user name. Clicking this hyperlink presents potential email addresses and social network profiles. In 2016, several new result categories were added including Documents, Web Search, Images, and Social Networks. The results here will likely have been discovered with the previous resources, but we can never have enough tools.

Quanki (quanki.com)

This service is extremely hit or miss. When your target does appear within this data set, you should receive full name, home address, email address, telephone number, and relatives. One unique feature is the presence of an entire date of birth. Unfortunately, I only locate my target on this service approximately half of the time.

WebMii (webmii.com)

This service emphasizes information associated with social networks. I have never located any home addresses or telephone numbers, but I have found online images that were not available on any other search engine. This is not the most productive option, but one to know when desperate for details.

Truth Finder (truthfinder.com)

There is nothing very special about this service, as it will likely have data similar to the other sites already mentioned. However, there is one major annoyance. When you search this site, you are bombarded by fake progress meters that insinuate that a huge report is being prepared about your target. On average, a real name search takes over 14 minutes due to these constant "please be patient while we find more details" notifications. The solution to this is to conduct the query from their information removal page at truthfinder.com/opt-out. This page was designed to identify personal records that you want removed, and the search is immediate. In my experience, you receive identical results as the traditional way.

Replicating explanations of additional people search databases will only display more redundancy than already present. I leave you with a few others that I have had success with in the past.

People Finder (peoplefinder.com)
People Finders (peoplefinders.com)
Check People (checkpeople.com)
White Pages (whitepages.com)
Research (research.com)
Speedy Hunt (speedyhunt.com)

IntelTechniques Person Search Tool (inteltechniques.com/osint/person.html)

The abundance of free person search tools can get overwhelming. They each have strengths and weaknesses, and none of them are consistent on accurate results. In years past, I would manually visit each site and enter my target information. This would usually result in small, yet valuable, pieces of information from each service. Today, I use a custom search tool that I created to search all possible sites simultaneously. This free tool will not locate any information that you could not find manually. Instead, it attempts to save you time by automating the process. Figure 10.01 displays the current state of this website. You can either enter the first and last name of your target within the search fields of each service; populate all fields by entering data at the first option; or enter this information only once at the final set of search fields. This latter option will launch several new tabs and conduct your search across each service. Note that browsers tend to block this type of activity unless you allow pop-ups for the domain.

The tool currently searches Pipl, That's Them, Spokeo, Advanced Background Check, Yasni, Intelius, Radaris, Zaba Search, Melissa Data, Fast People Search, Family Tree Now, True People Search, People Search Now, John Doe, WPNumbers, People Finder, People Finders, Truth Finder, Check People, White Pages, Quanki, PeekYou, WebMii, LinkedIn, and Twitter. The last two options replicate search queries discussed in previous chapters. During my live training courses, I am often questioned about two specific absences from this tool. The first is the inability to enter a middle initial or name. While I could make this an option, I find that a lot of people search websites do not always possess a middle name. Therefore, entering this data could harm an investigation by omitting valuable results. Some ask why I excluded some resources mentioned in the book. This is due to the way that those sites accept or refuse preset queries via a URL.

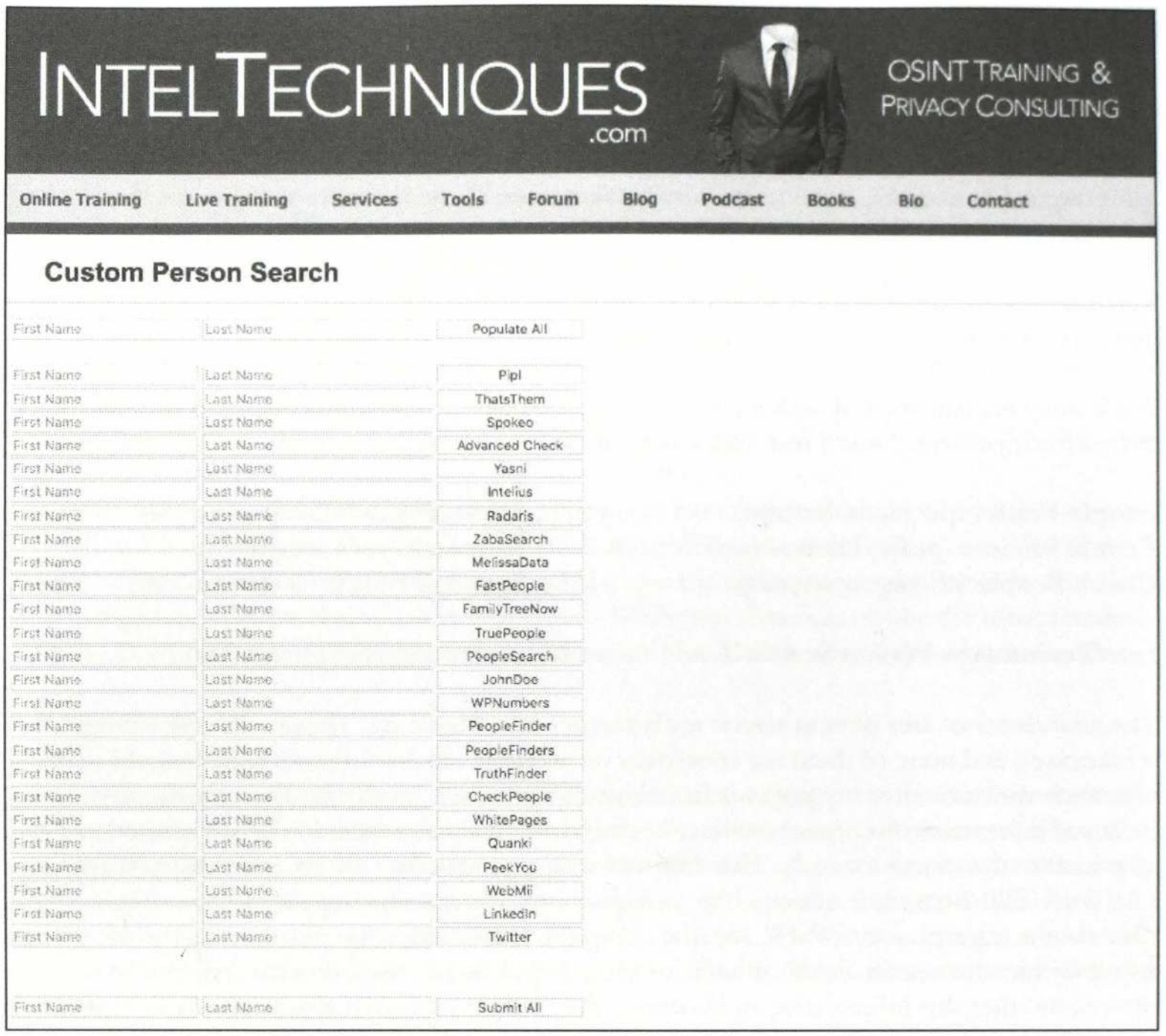


Figure 10.01: The IntelTechniques Custom Person Search Tool.

Putting It All Together

At first glance, the results obtained from free people search engines may appear redundant. There are definitely areas that repeat the same information. However, this can be beneficial for verification and building confidence in the results. Occasionally, what appears to be redundant might present further details. This can include a third report of an identical address that now includes an apartment number. In order to display the ways that each service can present unique details about your target, I conducted a query and documented all of the results. I chose the target of a close friend so that I could verify the accuracy of the results. I began my search with only his first and last name. The following details were identified which include the services that reported the information. Obviously, I have not included the actual data about my friend.

Age:	Pipl, Intelius, Radaris
DOB:	Yasni, WebMii
Current Address:	PeekYou, Spokeo, Yasni, Intelius, Melissa Data
Telephone Numbers:	Pipl, Radaris
Previous Addresses:	Radaris, Pipl, Quanki, White Pages
Previous Telephone:	Pipl, That's Them
Cellular Telephone:	True People, People Search Now
Spouse:	White Pages
Former Spouse:	Family Tree Now, Quanki
Mother:	Intelius, Pipl, Family Tree Now
Father:	Intelius, Pipl, Family Tree Now
Email Addresses:	Adv Background, Yasni
High School:	Yasni
College:	LinkedIn
Employer:	Intelius, Yasni
Occupation:	Intelius, That's Them
Facebook Profile:	Yasni
Twitter Profile:	PeekYou, Followerwonk
YouTube Channel:	PeekYou
Languages Spoken:	That's Them
Religion:	That's Them
Photos:	Pipl, WebMii

This summary was achieved in less than five minutes. Imagine what you could find if you took the time to continue analyzing social media links. The results above were all obtained from public data. These people search engines can assist greatly by aggregating this information.

How Many Of Me (howmanyofme.com)

This minimalistic site provides a simple interface to find out how many people exist that have a specific name. In my case, there are 16 people in the United States with my name. This is obtained from census data and can help determine how effective a targeted search will be. For example, if your target has a very unique name, you may still get numerous results to links of social network sites. In order to determine the likelihood that all of these profiles apply to the target, How Many Of Me can tell you whether your target is the only person that has that name. This site provides no intelligence about someone with a common name. I have used this in the past to determine appropriate covert names to use online.

Classmates (classmates.com)

Classmates is a very underrated resource for the internet searcher. Unfortunately, you must create a free account to take advantage of the worthy information inside the site. This free account can contain fictitious information and it is necessary to complete a profile on the site to access the

premium features. After you are logged in, you can search by first and last name. If you know the school that was attended, the results will be much more accurate. This should provide the school attended as well as the years the target attended the school. My new interest in this site is due to the availability of scanned yearbooks. The collection is far from complete, but there are a surprising number of complete yearbooks available to browse. This includes small towns and large cities. In one search, the Classmates embedded viewer displayed a page from the 1946 Alton High School yearbook. Analyzing this content can be very time consuming, as the yearbook must be manually browsed one page at a time. The information obtained should be unique from any internet search previously conducted.

Resumes

Resume searching was mentioned earlier in Chapter Three. Those methods will identify many documents, especially if the word "resume" is inside the file or file name. These results are only a portion of available content that could be extremely valuable to your investigation. I believe that resumes are an ideal target since they usually contain sensitive information that is not posted anywhere else. Many people will include a cellular number and personal email address on a resume, but would never consider placing these details on a social network. If the resume is publicly available, regardless of whether the target realizes this, we can gather good intelligence. The following techniques aim to assist in locating this valuable data. Detailed searches within Google or Bing will identify many resumes hosted publicly on websites and cloud-based document storage services. If my target's name is Michael Bazzell, I have found the following exact searches valuable on Google, Bing, and Yandex.

"Michael Bazzell" "Resume"
"Michael Bazzell" "Curriculum Vitae"
"Michael Bazzell" "CV"
"Michael Bazzell" "Resume" filetype:doc
"Michael Bazzell" "Curriculum Vitae" filetype:doc
"Michael Bazzell" "CV" filetype:doc
"Michael Bazzell" "Resume" filetype:pdf
"Michael Bazzell" "Curriculum Vitae" filetype:pdf
"Michael Bazzell" "CV" filetype:pdf
"Michael Bazzell" "Resume" site:docs.google.com
"Michael Bazzell" "Curriculum Vitae" site:docs.google.com
"Michael Bazzell" "CV" site:docs.google.com

While these queries will likely locate any resumes with text, they will fail on many resume images. Numerous resume hosting websites have realized that various data scraping engines scour their resume collection and "steal" their content. This has encouraged some services to store images of resumes that do not contain text that can be easily searched. While this is a decent layer of protection, it is not enough to keep out of Google results. Since Google scans images for Optical Character Recognition (OCR), it knows what words are within an image. After conducting the

above searches within traditional engines, attempt them within Google Images (images.google.com). A search of "Mary Johnson" "Resume" on Google Images revealed hundreds of images of resumes. A manual inspection of each identified many pieces of sensitive information.

CV Maker (cvmkr.com)

This website allows users to create free professional resumes and CVs. Currently, over 5 million have been created and are stored within the service. The home page does not offer a search option, as this service is not intended to be used as a people finder. However, we can rely on a Google search to get us the content we want. The following identifies the resume of our target.

site:cvmkr.com "john pratt"

The search result opens the resume page, which includes the following exact text.

E-mail: pr********@hotmail.com Phone: 0********** Address: * Staines**** *****, ****** ****

In the upper right corner, there is a button labeled "Download PDF". Within that file is the un-redacted content, which identifies his full email address, telephone number, and home address. On rare occasion, I have found this PDF option to be missing from my target profile. When this happens, we can create a direct link to the full details. In this example, our target's page is at cvmkr.com/7J0N. The following URL presents the entire PDF with the visible details. Basically, adding "?pdf=1" at the end of the URL should always present the full resume view.

https://cvmkr.com/7J0N?pdf=1

Since Google indexes all of the PDF files that are located, you can also perform searches for telephone numbers and email addresses using the site operator mentioned previously.

Indeed (indeed.com)

Indeed has a powerful collection of resume data. Because the term "resume" is not present in any of the content pages, you will likely not obtain this data during your standard searches. Entering your target name on Indeed under the "Find Resumes" option may present new results. Contact information is usually redacted. However, detailed work experience, education, and location are commonly present.

Ripoff Report (ripoffreport.com)

If your target conducts any type of business with the public, he or she will likely upset someone at some point. If your target regularly provides bad service or intentionally commits fraud within the business, there are likely many upset victims. Ripoff Report is a user-submitted collection of

complaints about businesses and individuals. I have had numerous investigations into shady people and businesses where these reviews by previously unknown victims were beneficial.

Gift Registries

Decades ago, people were surprised at the gifts presented to them after a wedding or birth. Today, we create online registries identifying the exact products desired, and within moments someone can purchase and ship the "thoughtful" gift with very little effort. As an investigator, I have always enjoyed the plethora of personal details within these registries, which tend to stay online long after the related event. Before identifying the best resources, let's take a look at the types of details we can acquire from some random targets.

Partner Name: When I am investigating someone, that person usually knows that they are under a microscope. He or she tends to stop posting to social media and start scrubbing any online details. However, their partner tends to ignore the threat of investigation and continues to upload sensitive information applicable to the target. Therefore, online wedding and baby registries help me identify the most lucrative target aside from the original suspect. In an example from the wedding registry website theknot.com, I received over 200 results for Michael Wilson, which also includes the name of the future spouse.

Maiden Name: In the example above, the results only identified future weddings. However, modifying the year in the search menu allows us to view past weddings. This will divulge a woman's maiden name. This can be beneficial in order to better locate a Facebook page or other family members that may be off my radar. I can also use this to search old yearbooks, criminal details, and previous addresses.

Date / State: Many counties will only share marriage certificates if the requestor knows the exact names of each party and the exact date of the event. We have everything we need in order to file a request. Marriage certificates often include full details of all parents, witnesses, and the officiant. Furthermore, I now have their anniversary date which can be helpful during a phishing attack or social engineering attempt. You might be surprised at the number of people that use their anniversary as a security question to an online account.

Ceremony Details: The Knot and other wedding registry sites offer the couple a free website to announce details about the upcoming (or past) event. This usually includes an embellished story about how they met, fell in love, and he proposed. While this could be good knowledge for social engineering, I am usually more interested in the wedding party. This will usually include the closest friends of my target, which will be next on my investigation list.

Items: While it may be fun to look at the items desired by a couple, there is much we can learn about their lives based on these details. In an example from Figure 10.02, I now know that a random Michael Wilson, who is getting married in San Antonio in November 2017, will be going to his honeymoon in Maui (#2), snorkeling (#3), at the airport carrying a Lowepro backpack

(#4), checking red/black suitcases (#5), capturing everything on a Canon HD camcorder (#6), dining at the Lahaina Grill (#7), and staying at a fancy nearby hotel (#8).

Other recent examples associated with actual targets identify the types of phones used, vehicles driven, and subjects of interest. While The Knot requires both a first name and last name to conduct a search, providing two asterisks (**) as the first name will present every entry online including the provided last name.

Children: The items within a baby registry will usually provide little to no value. Knowing the brand of diapers preferred or favorite crib style has never helped me in the past. However, knowing a due date and location of the target can be beneficial for future searching. Unfortunately, The Bump only allows searching of upcoming births, and not any past profiles. Fortunately, Google has our backs. The following Google search revealed multiple baby registries from the past few years associated with Michael Wilson:

site:registry.thebump.com "michael wilson"

Gifts: The most fruitful registries in regard to identifying personal preferences of a target are the various gift registries. Of all these, Amazon is the most popular. The following are the most common wedding, baby, and gift registries, with direct links to the most appropriate search pages. I highly encourage you to conduct a detailed Google "Site" search after attempting the proper method.

The Knot: https://www.theknot.com/registry/couplesearch
The Bump: https://registry.thebump.com/babyregistrysearch
Amazon Gifts: https://www.amazon.com/gp/registry/search
Amazon Baby: https://www.amazon.com/baby-reg/homepage/
Amazon Wedding: https://www.amazon.com/wedding/
Target Wedding: https://www.target.com/gift-registry/
Target Baby: https://www.target.com/gift-registry/baby-registry
Kohl's Wedding: https://www.kohls.com/gift-registry/wedding-registry.jsp
Registry Finder: https://www.registryfinder.com
My Registry: https://www.myregistry.com

Find a Grave (findagrave.com)

While I assume that your goal is to find living targets, you should also have a resource for locating proof of deceased individuals. I have used this website numerous times to locate the graves of recently deceased people. While not necessarily "proof" of death, it provides a great lead toward locating a death certificate and living family members.

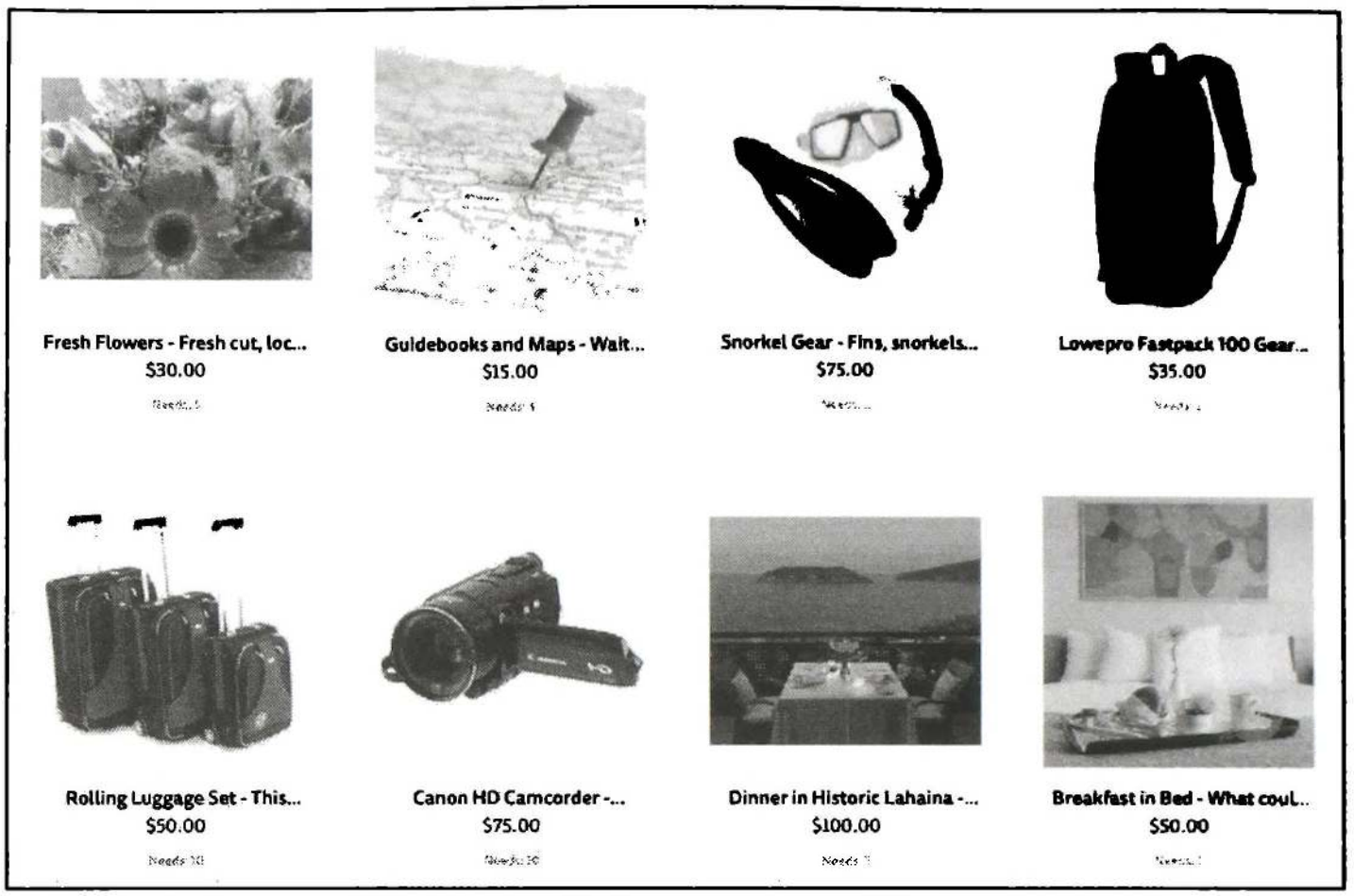


Figure 10.02: A search result from a gift registry website.

Addresses

The target of your investigation may be an address of your suspect. You may want to know who else lives at a residence. There are dozens of websites that possess databases of address information. I have outlined a few here that are unique from those already discussed. Additionally, the following websites which were previously discussed all allow reverse search of a residential address.

https://www.fastpeoplesearch.com/
https://www.whitepages.com
https://www.peoplefinder.com/reverse-address-lookup
https://www.peoplesearchnow.com/
https://www.truepeoplesearch.com/
https://radaris.com
https://www.intelius.com/property-records
https://www.advancedbackgroundchecks.com/address.aspx
https://www.spokeo.com/reverse-address-search
https://thatsthem.com/reverse-address-lookup
https://www.research.com/

White Pages (whitepages.com)

This is the official White Pages website that will conduct a reverse address search. Click on the "Address & Neighbors" tab and enter the address. Entering the zip code instead of the city and state will eliminate spelling errors or formatting inconsistencies. The results will include known residents and neighbors. Often, these neighbors listed will include current and previous residents. This data is pulled from public information and is rarely complete.

Voter Registration (www.blackbookonline.info/USA-Voter-Records.aspx)

Many people will have their address and telephone number unlisted in public telephone books. This prevents their information from appearing on some websites. If any of these people are registered voters, their address may still be public. In order to locate this data, you will need to connect to the county clerk of the county of residence. The link here will display a list of all fifty states. Clicking the state of the target will present all of the counties with known online databases of voter registration content. These will often display the full name of the voter and full address. This can be sorted by name or address depending on what information you have about the target. Chapter Eighteen presents additional voter registration search options.

Zillow (zillow.com)

This is a popular real estate information site. Entering an address will identify data such as purchase price history, sale status, satellite view map, estimated value, and surrounding real estate information. If the house is for sale or was recently for sale, and the home was listed on a real estate site, you will probably see the sale information here. If this includes interior photographs, which most do, you can view the interior of the house.

Google (google.com)

If all else fails, or you believe you are missing something, check Google. Searching the address should identify any leftover information about your target address. When searching, place the street address in quotes excluding the city. An example search may appear similar to "1234 Main" "Bethalto" IL. This will mandate that any results include the exact address and the exact city name, but they do not necessarily need to be right next to each other on the page. If you place the entire address including the city inside a single set of quotes, you would miss any hits that did not have this exact data. This search should reveal numerous home sale websites that may have unique interior photos.

Spokeo (spokeo.com)

Spokeo was explained earlier as a tool to search a target's real name. A reverse address search will also provide interesting information. Choosing the "Address" option and supplying a full address will identify the last names of any occupants. This is obtained from sources such as utility bills

and shipments. The first names of the subjects will be masked and only the first initial is visible. Paying a monthly fee will eliminate this masking, but is usually unnecessary. Instead, a custom search on Google will usually identify the target names. A site-limited Google search of your target's last name and address details usually identifies the full name of the target. The following search on Google would identify the full names of every person with the last name of Bazzell living at 121 Main Street in Houston.

site:spokeo.com "Bazzell" "121 Main" "Houston"

CHAPTER ELEVEN
TELEPHONE NUMBERS

There are hundreds of websites that claim the ability to search for information on telephone numbers and addresses. These vary from amazingly accurate results to sites that only include advertisements. If I have a target telephone number, there are three phases of my search. First, I want to identify the type of number and provider. The type could be landline, cellular or internet, and the provider could be the company supplying the service. Next, I want to identify any subscriber information such as the name and address associated with the account. Finally, I want to locate any online web content with a connection to the target number. This can all lead to more intelligence and additional searches. The majority of cellular numbers can now be identified if they are registered in someone's name. If you have an address, you will want to identify the people associated with the address and any telephone numbers the subjects use. This chapter will highlight the sites that can assist you with these tasks.

Carrier Identification

Ten years ago, I often queried telephone number porting websites to identify the provider of my target's telephone number. This would identify the cellular company that supplied service to my suspect. I would use that information within my court order to demand subscriber data about the target. Knowing the provider was essential as not to waste time requesting records from companies that had no data to provide. The websites used back then have either disappeared, or now charge a substantial fee for access. Five years ago, I noticed that an overwhelming amount of my target telephone numbers were connected to Voice Over Internet Protocol (VOIP) services such as Google Voice and Twilio. This would often be indicated by a result of "Broadband" or "Internet" instead of something obvious such as "Verizon". Until recently, the absence of a specific provider was a hurdle during investigations. Today, we have more sophisticated services that can identify exact provider details on practically any number.

Text Magic (textmagic.com/free-tools/carrier-lookup)

I begin with this option because it has been the most stable and does not limit searching as others do. This site requests any domestic or international telephone number and produces a report which includes the country, type of service, and provider associated with the number. While many online services can identify the provider of a cellular or landline number, this option provides the best detail about VOIP numbers. I submitted several telephone numbers associated with various providers that I could personally confirm. The following identifies the results of these searches. The first column represents the provider as I knew it, the second column is the type, and the third is the provider displayed by Text Magic. This is far from complete, but I wanted to demonstrate the ability to convert internet-based numbers into identifiable companies.

Verizon	Mobile	Verizon Wireless
AT&T	Mobile	AT&T Wireless
Google Voice	VOIP	Google/Level 3
Sudo (US)	VOIP	Twilio/Level 3 (SMS-Sybase) (MMS-SVR)
Sudo (Canada)	VOIP	Iristel Inc.
Sudo (UK)	VOIP	aql Wholesale
Sudo (FR)	VOIP	TRANSATEL
Blur	Landline	Twilio/Level 3 Communications
TextNow	VOIP	Enflick/Bandwidth.com (SVR)
On/Off	VOIP	Peerless Network

If this service cannot provide the data you need, or if you want another source to provide more confidence in the result, you should also consider **Carrier Lookup** (carrierlookup.com) and **Free Carrier Lookup** (freecarrierlookup.com). Carrier Lookup restricts users to one free search per day and Free Carrier Lookup has proven unreliable in the past. If all three of these services disappear by the time you need them, a Google search of "carrier lookup" should provide new alternatives. Once you have identified the provider, you should focus on the subscriber data associated with the number.

Caller ID Databases

In 2013, I began experimenting with reverse caller ID data. These are the same databases that identify a telephone number on your landline caller ID display. Often, this will include the name associated with the number. Until recently, this was something that only appeared on landline numbers, but that has changed. Now many cellular telephone numbers have name information associated with them. This name information is usually extracted from the cellular telephone service provider. I was immediately shocked at the accuracy of these results while searching cellular telephone numbers that were otherwise untraceable. On many of my investigations, this technique has eliminated the need to obtain court subpoenas to discover subscriber information.

The reason we can access this data is because it is necessary for telephone systems that do not already possess caller ID options. New systems that operate over the internet, referred to as VOIP systems (Voice Over Internet Protocol), do not receive caller ID data natively. This is something that we have taken for granted while it was provided by our telephone companies. Today, many businesses must purchase access to this data from resellers. This presents us with an opportunity.

I scoured the internet for every business that provides bulk caller ID data to private companies. Some offer a free web site for testing; some require you to submit queries through their servers; and others make you register for a free trial. I have tested all of them and identified those that are easy to access and give the best results. First, I will focus only on easy and reliable ways to search an individual number through specific web addresses.

Twilio (twilio.com/lookup)

This company provides VOIP services to many apps, companies, and individuals. An extended feature of their internet-based phone service is the ability to identify incoming calls through caller ID. Fortunately for us, they provide a page on their site that allows queries against their database, and use appears to be unlimited. On this page, replace the placeholder number (415-701-2311) with the number of your target. Allow the page to generate a result, which will appear directly below. The result will include the type of provider and any name associated with the billing. I have found this works best on landline and cellular telephone numbers and never responds with useable data from VOIP numbers. If I only had one website to search, this would be it.

Open CNAM (opencnam.com)

If I could only search two websites, this would be my second pick. Similar to Twilio, enter your target phone number and retrieve the results. You will usually receive the carrier and name associated with cellular and landline numbers. Open CNAM is one of the leading caller ID providers, and this page is designated for companies wanting to test the accuracy of their data.

WhoCalld (whocalld.com)

Similar to Open CNAM, this service will allow a single telephone number lookup within its caller ID database. It will also attempt to identify the current and previous cellular service carrier, as well as display the name associated with the number. The majority of my searches revealed redundant information from other services, but this can be valuable in order to increase your confidence in the results.

Caller ID Service (calleridservice.com)

The previous websites provided the easiest search options, but they might restrict you on the number of daily lookups. You might want to consider a more sophisticated approach that could provide unique data. This will require a bit more work on your part, but it will be justified. Caller ID Service has provided good accuracy with cellular telephone searches. You must register for the service to gain access to a free trial, and the process is very easy. Navigate to this website and register for a free account. Upon completion, you will receive an email with an API license key that is valid for approximately 20 free successful searches. You will not be charged for empty results. You must validate this address by clicking the link included in the message. This is their way of verifying that you provided an accurate and real email address. The following information was sent to my account.

User name: jwilson555
Password: mb555555
Auth KEY: 0b253c059b9f26e588ab101f4c2332b496e5bf95
Balance : 0.12

You are now ready to submit requests for caller ID information. To do this, you must formulate an API request that includes your user name, authentication key, and target telephone number to search. This is easier than it sounds. All we need is a number to search.

cnam.calleridservice.com/query?u=jwilson555&k=c2332b496e5bf95&n=6187271233

This queries the domain (calleridservice.com), our user name (jwilson555), our authentication key (c2332b496e5bf95), and our target number (6187271233). The service confirmed that this cellular number belongs to Craig Williams. I recommend saving the address of your first query as a bookmark or favorite. However, you should leave off the target telephone number at the end of the address. This will prevent the service from charging you for a credit every time you load this template. You can then add the new target number of interest at the end of the bookmark and conduct your searches easily. Caller ID Services grants you $0.12 in free searches, which will allow you up to 25 queries. Obtaining an additional free trial will only require a different email address.

Service Objects (serviceobjects.com/products/phone/reverse-phone-lookup-service)

In 2014, Service Objects removed their free online telephone lookup demo titled GeoPhone Plus 2. However, they will still allow you to generate a free API license which will allow you 500 free searches. Navigate to this website and complete the "Free API Trial Key" offer. You will receive an email similar to the following.

This is your DOTS GeoPhone Plus 2 API Trial License Key: WS77-OAZ3-xXxX

You can now use this key within a custom URL to search the registered owners of landline and cellular telephone numbers. The exact format is the following. Note that you would change "8475551212" to the target telephone number and "WS77-OAZ3-xXxX" to your trial license key.

http://trial.serviceobjects.com/gppl2/api.svc/GetPhoneInfo?PhoneNumber=8475551212&TestType=full&LicenseKey=WS77-OAZ3-xXxX

The response will be in XML data format. However, it will be easy to read. Below is an example.

<Provider><Name>NEW CINGULAR WIRELESS PCS, LLC - IL</Name>
<City>NORTHBROOK</City>
<State>ILLINOIS</State>
<LineType>WIRELESS</LineType>
<Name>JOHN ADORJAN</Name>
<Address>12142 S. 22nd<Address/>
<City>Chicago<City/>
<State>IL <State/>
<DateFirstSeen>2014-06-20</DateFirstSeen>

This entry identifies the target number as a wireless service provided by Cingular since June of 2014. The registered owner of the number is John Adorjan residing at 12142 S. 22nd in Chicago. Before reverse caller ID lookups, this information would have required a subpoena. In one recent example, a cellular telephone number searched on Service Objects revealed the name "Jennifer S" in the result. During an interview with this subject, she disclosed that "Jennifer S" is how she identifies her account on the telephone bill that she shares with other family members. She was unaware that this data was sent to the receiving number. On many searches, the full name will be present. This should explain why you may be noticing a caller's name on your caller ID display when he or she is calling from a cellular number.

Bulk Solutions (bulkcnam.com)

Bulk CNAM works very similar to Caller ID Services. You must register for a free account, and you will be granted limited free searches. You must provide a valid email address during the registration and will be required to validate that address after you receive an email with your license key. The custom address (URL) that we need to create only requires your license key and the target number. A user name is not necessary. The format of my free trial key is as follows.

cnam.bulkcnam.com/?id=b03c6513f688f89ee3f&did=6187271233

This queries the domain (bulkcnam.com), my free trial license (b03c6513f688f89ee3f), and my target telephone number (6187271233). A premium subscription to Bulk Solutions costs $0.009 per successful query.

CID Name (cidname.com)

The lower right portion of this page offers a free trial including 100 telephone number searches. The format for our address to query with this service is as follows.

https://dip.cidname.com/6187271233/d3a6e863c&output=raw&reply=none

This queries the domain (cidname.com), our target number (6187271233), and our license key (d3a6e863c). This service did not identify the name. However, when a name is not available, it will provide the general area of the telephone registration if available. This is a great example of why one service is never enough. I recommend that you use all of these services every time.

Open CNAM (opencnam.com/register)

Open CNAM was the second search option that was discussed in this chapter. Similar to the other premium caller ID search methods, this service also offers a free trial. You will receive a user name and access token. The following address contains the structure of a proper query.

http://api.opencnam.com/v2/phone/+16187271233?account_sid=f10&auth_token=AU5c43d8

This queries the domain (opencnam.com), the target number (16187271233), our account ID (f10), and our license number (AU5c43d8). Note that Open CNAM requires a "1" before the ten-digit number.

Everyone API (everyoneapi.com)

This service is owned by the same company as Open CNAM (Telephone Research LLC). The difference with this option is that it will display the cellular company that previously owned the number before it was ported. It will also search the number in a social network database. My research indicates that it is simply pulling Facebook data in the same way that was discussed in Chapter Four. This premium service is a bit more expensive than the previous options, but a complete telephone number search can be obtained during a free trial. The format of the URL for the request is as follows. You would replace "8475551212" with your target number, "xxx" with your account SID, and "yyy" with your license key. Below this URL is a typical response received. Note that Everyone API requires a "1" before the ten-digit number.

https://api.everyoneapi.com/v1/phone/+18475551212?data=name,carrier&account_sid=xxx&auth_token=yyy&pretty=true

"carrier": {"name": "Verizon Wireless"},
"carrier_o": {"name": "Cricket Wireless"},
"name": "Brian Williams"
"number": "+16189720000",

This result indicates that the target number was registered through Cricket Wireless before the current registration through Verizon Wireless. This number is currently associated with a social network profile of "Brian Williams".

Data 24-7 (data24-7.com)

This service provides $0.15 in free credit to anyone that creates an account. This balance provides approximately 40 free searches of their reverse caller ID service. The signup process is very similar to Caller ID Service. However, there is no email verification and you are provided user credentials without validation. You must choose a user name and password during the registration process which will be used within the URL submitted for each search. If your user name is jwilson and your password is lamp, the following address directly in your browser would identify subscriber information from the cell number of 6187271233. The result is below the demo.

https://api.data24-7.com/v/2.0?user= jwilson&pass=lamp&&api=I&p1=6187271233

<response>
<results>

<result item="1">
<status>OK</status>
<number>16187271233</number>
<name>Craig Williams</name>
</result>
</results>
</response>

Next Caller (nextcaller.com)

Next Caller provides a typical reverse caller ID service with one big difference. It also provides addresses for the subscribers of the landline and cellular target numbers. It will take a bit more effort to execute this service for our needs, but the work will be worth it. Navigating to this site allows you to create a free user account. It is useless until a representative activates your free trial. This will usually require a brief phone conversation about your needs. I do not recommend telling them that you only want to search numbers. Appearing to be a much bigger potential client will work well in your favor. After your account is activated, you can begin conducting queries. Lately, readers have informed me that they had to put a lot of effort into convincing a sales representative for a free trial. Of those that were successful, all were extremely happy with the data in the results.

The user portal will identify how to program this service into your server using various types of computer programming languages. They do not provide a URL solution similar to the previous examples. Instead, we will need to add a bit of technology. Next Caller accepts curl requests to their database and will return complete results. Curl is a command line tool for getting or sending files using URL syntax. Next Caller will provide you with a user name and password for your free trial account. For the purposes of this demonstration, assume that your user name is aaaaa and your password is bbbbb. Your target telephone number is 202-555-1212. From a command prompt, you could submit the following commands.

curl -X GET \
-u "aaaaa:bbbbb" \
-H "Content-Type: application/json" \
"https://api.nextcaller.com/v2.1/records/?phone=2025551212&format=json"

This can get quite tiresome when conducting multiple searches. Instead, type the following exact URL into the address bar of your web browser.

https://aaaaa:bbbbb@api.nextcaller.com/v2/records/?phone=2025551212&format=json

This sends a curl request including your user name (aaaaa) and password (bbbbb) to Next Caller (@api.nextcaller.com). It asks for a JSON style request for the target number. The response will be much longer than the previous services. The following is an actual example from a cellular number in Florida with redacted information. All fields were visible in my result.

first_name: "Stephen",
last_name: "REDACTED ",
name: "Stephen REDACTED",
language: "English",
phone: number: "618972XXXX",
address city: "Tampa",
line1: "28XX Falcon XXXX",
state: "FL",
zip_code: "XXX"
line_type: "Mobile",
carrier: "New Cingular Wireless Pcs, Llc",

The result provided the complete customer name, billing address, and cellular provider. Caller ID database results are my current favorite way to extract cellular and landline information from telephone numbers. They have been much more reliable than standard website search engines that often display inaccurate and dated information. Almost every day, I am contacted by an investigator that is stuck on a telephone number involved in an investigation. Lately, these are usually Craigslist style theft cases or subjects inappropriately contacting children over mobile devices. If the telephone number is registered to someone, we have a 90% success rate in identifying the person through these methods. If the telephone is a "burner" style device that is not registered to anyone, these methods will not produce any valid results.

There are other caller ID options available on the internet. I encourage you to investigate any companies that have surfaced since this research. Some services may no longer offer a free trial period to test the database. I have purchased premium memberships through CID Name and Open CNAM. A $10 purchase will allow you over 1,000 queries at each provider. In Chapter Twenty, I will explain how you can create a simple web page that will query all of these services at once when you provide a single telephone number. The results will all display immediately for you. I use my own custom page every day. If this is overkill for your needs, there are other web-based search engines that are easier to use.

Caller ID Test (calleridtest.com)

This site was designed to input a telephone number and test the Caller ID display feature. It is nothing more than a standard lookup service, but I have found the data to be unique from other sources on some occasions. Unfortunately, I have also found the availability of this service to be completely unreliable. While the site is usually present, the results don't always populate. However, this resource should be checked as a last resort when the other processes have failed.

Overall, reverse caller ID services can tell us more about a target telephone number than the standard people search engines. In many cases, you can immediately obtain data that would have required a subpoena just a few years prior. Always utilize all of the services in order to gauge the confidence in the results.

Pipl (pipl.com)

One of the easiest search engines for telephone numbers is Pipl. This is the same search page that was discussed for name searches earlier in the book. The same search field will handle a telephone number. This number should not be entered in the standard separated format such as 555-445-8543. Instead, enter the number without dashes or spaces similar to 5554458543. This can increase the accuracy and number of results. The results page will first identify any profiles created with the target number. These profiles can be visited for more content related to the number. The information received from this profile often includes full name, address, date of birth, and relatives. The "Suggested searches" area contains the profiles created by Pipl that often identify the year of birth, complete address, and several relatives. Scrolling further down this page would display links to more results that would verify this information.

This is not the limit of the information you can get from Pipl about a telephone number. There is nothing else to search on the official profile, but there is data stored that you may want. Pipl maintains a street listing of all of the households on a specific street that is cross-referenced by telephone number. Unfortunately, there is no simple way to access this on the Pipl site. To get this information, you can use Google. Conduct the following style of search on the Google main page, replacing the listed number with the target number.

site:pipl.com 5554458543

This should provide a single result that will link to the Pipl page that identifies all residences on the same street as the house with the target telephone number. This lists the telephone number and address of each household. The telephone numbers are links that will open a profile on that number.

Real World Application: As a detective, I was tasked on several occasions to assist with background checks on Police Officer applicants. Occasionally, the current address of the applicant is out of state and visiting the neighbors is not optimal. Entering the telephone number of the applicant's residence in this fashion will display all of the neighbors' telephone numbers and addresses. Visiting the associated profiles will identify enough information to make contact via telephone and conduct an interview.

Fast People Search (fastpeoplesearch.com)

This service was also mentioned in the people search chapter, but their telephone search is equally important. Selecting the "Phone" option from the main page and entering your target number should be sufficient. This can also be executed from a direct URL as follows.

https://www.fastpeoplesearch.com/618-462-0000

True People Search (truepeoplesearch.com)

The name search databases of this option and the previous appear identical. However, I occasionally receive additional telephone results from this site. Queries can also be submitted via the following URL structure.

https://www.truepeoplesearch.com/results?phoneno=(618)462-0000

Facebook (facebook.com)

In Chapter Four, a method of searching cellular telephone numbers on Facebook was explained. This technique is currently one of the most successful methods of identifying the owner of a cellular number. If there is any chance that your target is on Facebook, the number should be searched with that technique.

Additional Resources

Explaining how to enter a telephone number into every reverse phone search website is unnecessary. Instead, I present the additional resources that I have found helpful in my investigations. They are listed in order of most benefit to least.

Advanced Background Checks (advancedbackgroundchecks.com/phone)
Nuwber (nuwber.com/phone)
Sync.me (sync.me)
US Phonebook (usphonebook.com)
White Pages Plus (whitepages.plus)
Search Bug (searchbug.com/peoplefinder/phone-search.aspx)
OK Caller (okcaller.com)
That's Them (thatsthem.com)
411 (411.com)
Who Calls Me (whocallsme.com)
800 Notes (800notes.com)
Number Guru (numberguru.com)
Reverse Genie (reversegenie.com/reverse_phone)
Super Pages (wp.superpages.com)
Yahoo (people.yahoo.com)
Free Phone Tracer (freephonetracer.com)
Fone Finder (fonefinder.net)
Mobile Phone No (mobilephoneno.com)
Skip Ease (skipease.com/reverse)
IvyCall (ivycall.com)
Numpi (numpi.com)

Burner (challenge.burnerapp.com)

Burner is a VOIP app that provides a disposable number to use in conjunction with your cellular telephone. In order to generate new interest in their product, they have created an online tool that identifies sensitive information associated with a telephone number. Much of the data provided appears to come from the same sources mentioned previously. However, it works better than any other resource when searching VOIP numbers. I conducted a query of one of my Google Voice numbers and received the following information.

Mike Bazz***
1700 E. B***dway, *****, IL

While the information is heavily redacted, I can logically assume that the data matches my real information. The address is my previous workplace. It should be noted that I used this number in a very overt capacity and associated it with my real name on a few websites. However, I could not reproduce these details using any other service. To use this free service, simply enter any target number. The results will disappear from the screen quickly, so prepare a screen capture tool before the search.

Live Escort Reviews (liveescortreviews.co)

If you have any suspicion that the target of your investigation is involved in prostitution, drugs, or any related activity, Live Escort Reviews should be checked against the telephone number of your subject. This website aggregates all of the prostitution classifieds and review sites into one search. It extracts the telephone numbers from all online classified pages and allows you to search by the target telephone number. One of my training examples identified 37 online photos, 20 escort ads, several reviews by "Johns", ages used by the target, the last known location, and locations visited based on postings in online classifieds. Any time I have a target telephone number that is likely involved in criminal activity, I conduct a brief search on this site.

Search Engines

Google and Bing were once a great place to find basic information about a target phone number. These sites can still provide valuable information, but the amount of spam that will display in the results is overwhelming. Many of the links presented will link to sites that will charge a fee for any information associated. This information is usually the same content that could have been located with an appropriate free search. I do not recommend giving in to these traps. While we can't ignore a traditional search of telephone numbers, we can customize the queries in order to achieve the best results. Before explaining advanced telephone owner identification, we should take a look at appropriate search engine structure.

Most people use traditional search engines as a first stop toward identifying the owner of a telephone number. The number is usually provided in a standard format such as 202-555-1212.

This can confuse some search engines because a hyphen (-) is often recognized as an operator to exclude data. Some engines might view that query as a search for 202 but not 555 or 1212. Additionally, this search might identify a website that possesses 202-555-1212 within the content but not one that contains (202) 555.1212. If this is your target number, all of the following should be searched in order to exhaust all possibilities. The quotation marks are important to prevent the hyphen from being seen as an operator.

"2025551212"
"202-555-1212"
"202.555.1212"
"202 555 1212"
"(202) 5551212"
"(202) 555-1212"
"(202) 555.1212"
"(202)5551212"
"(202)555-1212"
"(202)555.1212"

This may seem ridiculous, but I am not done. Many websites forbid users to post a telephone number, such as many auction sites, but people try to trick this restriction. They will type out a portion of their number to disclose contact information. While not a complete list of options, the following should also be searched.

"two zero two five five five one two one two"
"two zero two five five five 1212"
"two zero two 555 one two one two"
"two zero two 555 1212"
"202 five five five one two one two"
"202 555 one two one two"
"202 five five five 1212"

This list would not capture a post that included (202) 555 twelve twelve, but you get the point. After submitting these through Google, you should attempt each through Bing. In my effort to always provide online tools that automate and simplify these techniques, I have created a telephone search tool at the following website.

https://inteltechniques.com/intel/OSINT/telephone.html

The lower portion of this page, displayed later in Figure 11.01, allows you to enter a numerical and written target telephone number. Clicking the submit button launches a series of JavaScript commands that launch eight new tabs within your browser. The first four are custom Google searches with the target data and the last four repeat the process on Bing. The following four searches are conducted on both services, using the example data entered previously.

"2025551212"OR"202-555-1212"OR"202.555.1212"OR"202 555 1212"

"(202) 5551212"OR"(202) 555-1212"OR"(202) 555.1212"OR"(202)5551212"OR"(202)555-1212"OR"(202)555.1212"

"two zero two five five five one two one two"OR"two zero two five five five 1212"OR"two zero two 555 one two one two"OR"two zero two 555 1212"

"202 five five five one two one two"OR"202 five five five one two one two"OR"202 five five five 1212"

Notice that these queries use quotation marks to obtain exact results and the OR operator to search multiple options independently from each other. You will likely receive many false positives with this method, but you are less likely to miss any relevant results. While this is a great starting point for number searches, it is much less reliable than the next method.

True Caller (truecaller.com)

This service stands alone as the most creative telephone number lookup service. True Caller is an app for smart devices that displays caller ID information of incoming calls. If you receive a call on your phone, and the number is not in your contacts, True Caller searches its database and provides any results on your screen. You can then choose to accept or deny the call. This is fairly standard and is not the interesting aspect of this service. The fascinating part to me is the source of their caller database. It is completely crowd sourced. When you install the app, you give it permission to collect all of your contacts and upload them to the master database. Basically, millions of users have uploaded their contact lists for the world to see. The next amazing thing to me is the ability to search within this data on the True Caller website. You must connect to the service via a covert Facebook or Yahoo account, but that is not difficult. When I first found this service, I was skeptical. I entered the cellular number of my government issued cellular telephone expecting to see no results. The response was "Mike Bazell". My jaw dropped. My super-secret number was visible to the world. This means that someone in my circle, likely another government employee, installed True Caller on his or her phone and had my information in their contacts. Until someone else populates data for this number, it will always be present in the database.

IntelTechniques Telephone Search Tool (inteltechniques.com/intel/OSINT/telephone.html)

As mentioned earlier, I have created a custom search tool to assist with complete queries for Google and Bing. Additionally, this tool automates several searches through third-party websites. Figure 11.01 displays the current state of the tool. The first section provides a single area to enter the target telephone number and populate all remaining fields. You can then submit this number through Facebook, True People Search, Fast People Search, Advanced Background Checks, Pipl, 411, US Phonebook, White Pages Plus, That's Them, True Caller, Sync.me, Who Calls Me, ZabaSearch, Dex Knows, Burner Challenge, OK Caller, Search Bug, and Live Escort Reviews. In my experience, all of these services can be queried in less than five seconds. Each result will open in a separate tab in your browser for easy analysis.

The right column offers audio recordings of the stock voicemail greetings for the major U.S. carriers. These can be used for comparison when listening to your target's stock greeting. These will likely identify the service provider in the event the online search tools fail.

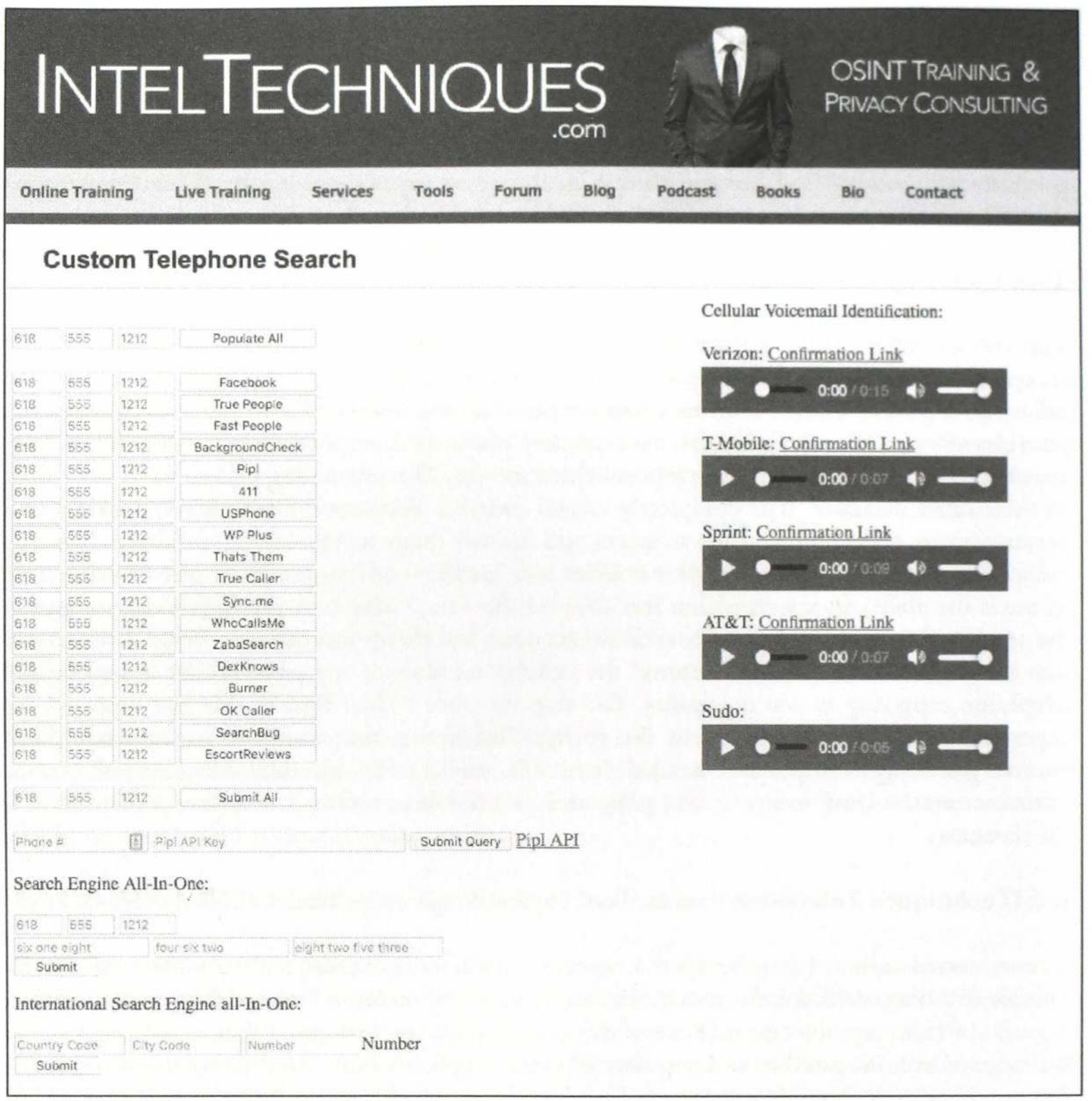


Figure 11.01: The IntelTechniques Custom Telephone Search Tool.

Craigslist (craigslist.org)

Craigslist has already been discussed in earlier chapters, but the phone search options should be further detailed. Many people use Craigslist to sell items or services. The posts that announce the item or service available will often include a telephone number. These numbers will belong to a landline or cellular provider. This can be a great way to identify unknown telephone numbers.

Some posts on Craigslist will not allow a telephone number to be displayed on a post. It is a violation of the rules on certain types of posts. Some people choose not to list a number because of automated "scrapers" that will grab the number and add it to databases to receive spam via text messages. Either way, the solution that most users apply to bypass this hindrance is to spell out the phone number. Instead of typing "314-555-1212", the user may enter "three one four five five five one two one two". Some will get creative and post "314 five five five 1212". This is enough to confuse both Craigslist's servers as well as the spammers. This can make searching difficult for an analyst. The hard way to do this is to conduct several searches similar to the following.

site:craigslist.org "314-555-1212"
site:craigslist.org "314" "555" "1212"
site:craigslist.org "three one four" "five five five" "one two one two"
site:craigslist.org "314" "five five five" "1212

This list can get quite long if you try to search every possible search format. One search that will cover most of these searches in a single search attempt would look like the following.
site:craigslist.org "314"|"three one four" "555"|"five five five" "1212"|"one two one two"

The "|" symbol in this search is the same as telling Google "or". In essence, we are telling Google to search "314" or "three one four", then "555" or "five five five", and then "1212" or "one two one two". With this search, you would receive a result if any combination of the following was used.

314-555-1212
314.555.1212
3145551212
314 555 one two one two
three one four 555-1212

This search will not catch every possible way to post a phone number. For example, if the user had typed "314 555 twelve twelve", the above technique would not work. The researcher must consider the alternative ways that a target will post a number on a website. It may help to imagine how you would post the target number creatively on a site, and then search for that method. Additionally, searching for only a portion of the number may provide results. You may want to try searching only the last four digits of the number. This may produce many unwanted results, but your target may be within the "haystack". This technique is not unique to Craigslist only. The same searches would get you results on other sites, such as Backpage, by changing the domain name as follows.

site:backpage.com "314"|"three one four" "555"|"five five five" "1212"|"one two one two"

Spy Dialer (spydialer.com)

This service takes a new approach on identifying the user of a cellular telephone. While it offers a typical telephone number search tool, which appears to extract data from crowd-sourced databases, the real power lies in the voicemail retrieval. Most cellular users have an outgoing voicemail message that identifies them by name. Others create a custom message with their own voice. Any of this can help determine the user of the number. Spy Dialer attempts to connect to the service provider of the cellular number; extract the outgoing voicemail message; and present it to you in mp3 format for listening and downloading. All of this is normally completed without ringing the target's telephone. You should note that on some occasions, the target telephone number rang during testing. It did not identify the caller, but the unusual single ring may raise suspicion with a paranoid target. In my experience, this happens about 10 percent of the time. If the target dials the Nevada-based number of the missed call, he or she will be notified that a Spy Dialer call was placed. My successes with this method outweigh the risk. I have had several investigations that involved "anonymous" cash cellular telephones that announced the owner's name on the outgoing message. When a successful result is displayed, you can click the link below the player to download the audio file to your computer.

Sly Dial (slydial.com)

This service conducts an inquiry into a cellular telephone the same way as Spy Dialer's attempt. It contacts the cellular provider of the telephone number and sends you straight to the outgoing voicemail message of the target. However, there are two big differences.

Sly Dial does not work through a website. Instead, you must call a general Sly Dial telephone number and follow the automated prompts. You must listen to a brief advertisement before your call is placed. Finally, the service will play the target's outgoing voicemail message through this audible telephone call. Since a website is not involved, there is no option to download an audio file of the call. We can obtain an audio copy of this message by placing the call through Google Voice and recording the session by pressing the "4" button on the dial pad.

Sly Dial does not usually ring the suspect's telephone. It will likely not show "missed call" or any other indicator that a call occurred. In my testing, less that 5 percent of the attempts actually cause the target telephone to ring only one time. Calling the missed call back reveals nothing about the identity of the number. Ultimately, there is a very small chance that the target will know that someone attempted a call. In the rare occurrence that the telephone rings, the target will never know the identity of the person making the calls. To use the Sly Dial service, call 267-759-3425 (267-SLYDIAL) from any telephone service including landlines, cellular lines, or VOIP. Follow the directions during the call. If this number does not work, visit slydial.com for updates.

I want to stress the following one additional time. Use these services at your own risk. If accidentally notifying your target that you are conducting these types of activities could compromise your investigation, avoid these two techniques.

Infobel (infobel.com)

Infobel can conduct telephone and address searches on international subjects from almost every country. Several of the links will forward you to external third-party services that will appear within an Infobel window. Any links identified as "Infobel" will search an internal database of numbers. The results will identify telephone numbers and addresses for the target searched.

Grocery Reward Cards / Loyalty Cards

Most grocery chains have adopted a reward/loyalty card system that mandates the participant enroll in their program. The consumer completes an application and receives a plastic card to use during checkout for discounts. Many of these stores only offer a sale price if you are a member in the program. Most consumers provide a cellular telephone number to the program and use that number during checkout. This eliminates the need of possessing a physical card in order to receive the discount. Instead, they type their cell number into the card swiping machine to associate the purchase with their membership. These programs contain a huge database of telephone numbers and the registered users. There is no online database to access this data. However, you can obtain this data if you are creative.

Assume that your target telephone number is 847-867-5309. If you have tried every technique mentioned at this point to identify the owner and failed, you may consider a query with a local grocery chain. The easiest method is to enter the store, purchase a pack of gum, and enter the target telephone number as the reward/loyalty program number. You will likely receive a receipt with the target's name on the bottom. Figure 11.02 (left) displays a portion of the actual receipt that I received when using this number. If you prefer to avoid entering a store, drive to the company's gas station outside of the store. Figure 11.02 (right) displays the notification I received when entering this same number at the pump. Note that this number is fictional. However, it has been registered at practically every grocery store in the United States. Try to use it the next time you make a purchase.

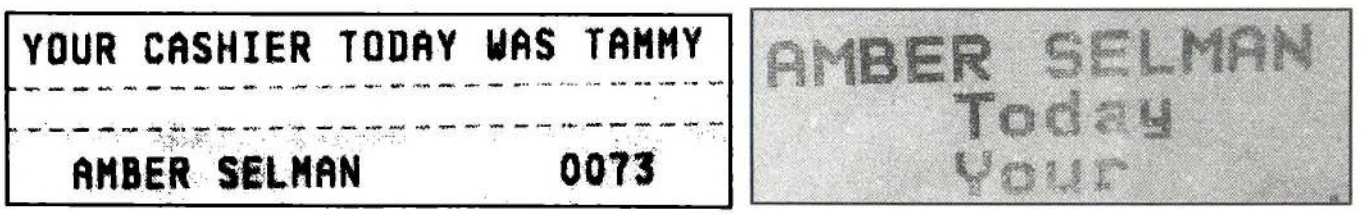


Figure 11.02: A receipt (left) and gas pump (right) identifying the owner of a cell number.

Chapter Twelve
Online Maps

The presence of online satellite images is not news anymore. Most of you have already "Googled" your own address and viewed your home from the sky. This view can get surprisingly detailed when using the zoom feature. Alleys, sheds, and extended driveways that are hidden from the street are now visible thanks to this free service. Many tactical units will examine this data before executing a search warrant at a residence. Aerial maps are helpful for viewing the location of exiting doors, escape routes, stairs, and various obstructions. The rapidly growing availability of the Street View option now gives us more data. This chapter explains detailed use of Google, Bing, and other mapping services. At the end, I present my custom maps tool which provides an automated solution to collecting every possible view associated with your target of interest.

Google Maps (maps.google.com)

In 2014, Google made several changes to their online maps service. They introduced a new feature with Street View that made the default view full screen. This eliminates the Google search bar, side menu, browser menus, and any other items from blocking a larger view. Additionally, Google streamlined the entire Maps experience to make everything easier to use. Unfortunately, they also eliminated many of the features that were beneficial to researchers and investigators. Fortunately, they have re-enabled some of these missing features. The following basics of Google Maps are now default for all users.

Search Bar: The Google Maps search bar can now accept practically any type of input. A full address, partial address, or GPS coordinates will immediately present you with a mapped view. Company names and types of businesses, such as "café" will highlight locations that may be of interest. This search field is the first stop. Attempt any search relevant to your investigation and you may be surprised at how accurate Google is. You can collapse this entire menu by clicking the left arrow next to the search field. This will return you to a full screen view of the map.

Satellite/Earth View: The lower left area of any map will offer a satellite and earth view. The satellite view is a direct view from the sky looking almost straight down. The Earth view is similar, but offers the tilt option. While in the earth view, click on the small icon to the right of the map that appears similar to four small squares. This will shift the view 45 degrees and a second click will shift an additional 45 degrees. A third click returns to the standard satellite view. The rotation icon above this button allows you to rotate your view for the desired result. While satellite views of maps are now well-known, we see continuous enhancements that are not advertised. A satellite view of your target location is always vital to every investigation.

Street View: If the Street View option is available, Google has been in the area and captured a photo of the location from the street. Dragging and dropping the small orange man in the lower right menu will open a street view from ground level at the area specified. You can navigate through this view by clicking forward, clicking and dragging, or scrolling and zooming. This view can be zoomed in by double-clicking and panned left and right by dragging the mouse while holding a left click. Double-clicking an area of the street will refresh the window to the view from that location. Clicking "Back to map" in the lower left will return you to the standard map view.

Historic Street View: In late 2014, Google began offering the ability to view all stored street view images for any single location. This option is available within the standard street view layout within the search area of the upper left corner. Click on the small clock in order to launch a pop-up window. This new view will allow you to move a slider bar which will present different views. The month and year of image capture will also appear for documentation. Figure 12.01 displays this method which presents an additional view of a parking lot from a few years prior. Additional options include views from several years prior. This can often reveal additional vehicles or missing structures associated with an investigation.

Distance Measurement: Google Maps reintroduced the distance measurement tool after completely disabling the classic maps interface in 2015. While in map or satellite view, right-click on your starting point and choose "Measure distance". Click anywhere on the map to create a path you want to measure. Further clicks add additional measuring points. You can also drag a point to move it, or click a point to remove it. The total distance in both miles (mi) or kilometers (km) will appear under the search box. When finished, right-click on the map and select "Clear measurement".

GPS Coordinates: Clicking on any point will load a small window in the bottom center that identifies the exact GPS coordinates of the chosen location. If this is not visible, right-click any point and select "What's here".



Figure 12.01: Historic Street View options from Google Maps.

Bing Maps (bing.com/maps)

Similar to Google Maps, Bing offers a map view, satellite view, and street view. Bing does offer something that is not always available in Google. Bing possesses a "Bird's Eye View" which displays four distinct angled views of a location. This may display signs, advertisements, pedestrians, and other objects with clear visibility. While Google is rolling out their own 45-degree satellite view, the areas covered are minimal at the time of this writing. In my experience, the imagery provided by Bing is often superior in quality compared to Google Maps. A side by side comparison can be seen in a few pages with the custom maps tool. By default, this view will always be of the south side of a location, looking north. The curved arrows in the upper right corner allow you to navigate to three additional views which display the west, north, and east sides of the location.

Dual Maps (data.mashedworld.com/dualmaps/map.htm)

This website provides a satellite view of a location on both Google Maps and Bing Maps simultaneously. The Google view on the left will also contain a search field in the lower left corner. Searching an address in this field will center both maps on the same address. This will provide a comparison of the satellite images stored on each service. This can quickly identify the service that has the better imagery. While this can provide a quick side-by-side comparison, upcoming automated solutions are preferred.

Here Maps (here.com)

Another option for an alternative satellite view of a location is Here Maps. The areas that display a detailed view are limited, but worth investigating. If you are searching for imagery of a large city, the 3D view displays great detail of the buildings and structures. This imagery is usually independent of Google and Bing data, but some areas incorporate content from Microsoft.

Zoom Earth (zoomearth.com)

This multiple satellite imagery website presents views from NASA, Bing, and ArcGIS. Occasionally, the ArcGIS data is more recent than Google or Bing. The smooth interface will easily provide a comparison of the available images for any location. One advantage of Zoom Earth is the ability to view satellite images in true full-screen mode. This allows creation of full-screen captures without branding, menus, or borders. This could be more appropriate for live demonstration instead of a standard Google or Bing window.

Map Box (mapbox.com)

While these satellite images only offer city-level views, you will not find a more recent collection of data. At the time of this writing, 12/27/17, the satellite view of Chicago was collected on

12/20/17 (seven days prior). Additionally, the site notified me that the next collection of images is set to occur on 1/5/18 (in nine days).

Descartes Labs (descarteslabs.com)

This service offers a very unique search option that I have not found present on any other site. After locating a target of interest, it displays a satellite view sourced from the National Agriculture Imagery Program (NAIP). The unique part is the ability to search based on image. In other words, you can select a monument, park, building, or any other view and request any images that match. As an example, I selected the baseball grounds at Wrigley Field and I was immediately presented hundreds of baseball fields all over the world. I have yet to determine how I would execute this strategy within an investigation, but this feature has potential.

Crowd-Sourced Street Views

Satellite and Street View maps from services such as Google and Bing are nothing new. Most of you can view the top and front of your home from multiple online websites. With street view options, these services are fairly responsible and block most faces and license plates. This makes it difficult for investigators trying to identify a suspect vehicle parked at a home or present at a crime scene months prior to the incident. We have two services that offer unique street-level views that may remove these limitations. **Mapillary** (mapillary.com) and **Open Street Cam** (openstreetcam.org) appear similar to other mapping websites when first loading. You see a typical map view identifying streets, landmarks, and buildings. Enabling the satellite view layer displays images from the Open Street Map project. The real power is within the crowd-sourced street view images. While in any map view, colored lines indicate that an individual has provided street-level images to Mapillary or Open Street Cam, usually from a GPS-enabled smart phone. This is actually quite common, as many people record video of their driving trips that could be used in case of an accident. These services make it easy and automated to upload these images. The sites then embed these images within their own mapping layers for the public to see.

Clicking these colored lines reveals the street view images in the lower portion of the screen. Expanding these images allows you to navigate through that individual's images similar to the Google Street View experience. Figure 12.02 displays a street view layered over a satellite view. The user name of the Mapillary member and date of image capture appears in the lower left.

In some of these images, the services appear to be redacting license plates with a typical "Blur Box" as seen in Figure 12.03 (left). A few feet later, the box disappears and a partially legible license plate is revealed, as seen in Figure 12.03 (right). It seems like these services are attempting to determine when a license plate is legible, and then blurring it. When the plate is farther away and more difficult to read, it is ignored. This can work in our favor. In Figure 12.04 we can use selective cropping and photo manipulation to obtain the registration. The left image appears unaltered as it is difficult to read. The right image was cropped; inverted with Photoshop;

brightness turned down to 0%; and contrast heightened to 100%. The result is a legible license plate. I believe most registration plates can be made visible within these services.



Figure 12.02: A crowd-sourced street view from Mapillary.



Figure 12.03: A vehicle with a blurred registration plate (left) and clear (right).



Figure 12.04: An illegible registration plate (left) and manipulated view (right).

These sites allow you to identify the uploader's user name, mapping history, number of posted images, and profile image. You can also select to watch all of the captured images from a specific user as he or she travels daily. I can't begin to imagine the amount information available about a user's travel habits if he or she were to become a target of an investigation. While there is not coverage of every area like we see with Google Maps, the databases are growing rapidly, and should be included when using other mapping tools.

Historic Imagery

Researching different satellite views of a single location can have many benefits. These views are all of the current content stored within each service. However, these mapping services continuously update their offerings and usually present the most recent option. You may want to view the previous content that was available before an image was updated. Chapter Nineteen will explain software options for retrieving older images. Additionally, some web-based services also offer alternative views.

Historic Aerials (historicaerials.com)

If you need satellite imagery from several years prior, you can visit Historic Aerials. The quality will often be poor, especially as you view imagery from previous decades. After you enter an address, you will be presented all available options on the left side of the page. Figure 12.05 displays several results of the same location over a twenty-year period. These views will be unique from all of the previously mentioned services.

Terra Server (terraserver.com)

This repository of satellite images dating back to 1997 offers an impressive alternative view of most locations on earth. After selecting your location of interest by name, address, or GPS coordinates, you can zoom to levels comparable to Google Maps. A free account is required for zoom levels that display homes. The historic satellite images were provided by Sovinformsputnik (the Russian Federal Space Agency) and GeoEye. A search of Wrigley Field in Chicago produced twelve unique views from 2012 through 2017.

Land Viewer (eos.com)

This resource will not present detailed views of your target's home. The images here are often generated from weather satellites, and restrict zoom levels to a city view. Most locations offer four active satellites that constantly retrieve images and five inoperative satellites that store historic imagery dating back to 1982. I have only used this resource to document potential weather at a crime scene (clouds, rain, snow, or clear).

Real World Application: Combining several satellite views can provide much information about a target's residence. Before the execution of a search warrant, it is beneficial for police to collect

as much map information as possible. This will give updated views of a drawn map, satellite imagery directly above the house, four angled views from the sky, and a complete view of the house and neighboring houses from the street, including vehicles. This can be used to identify potential threats such as physical barriers, escape routes, and video surveillance systems in place.

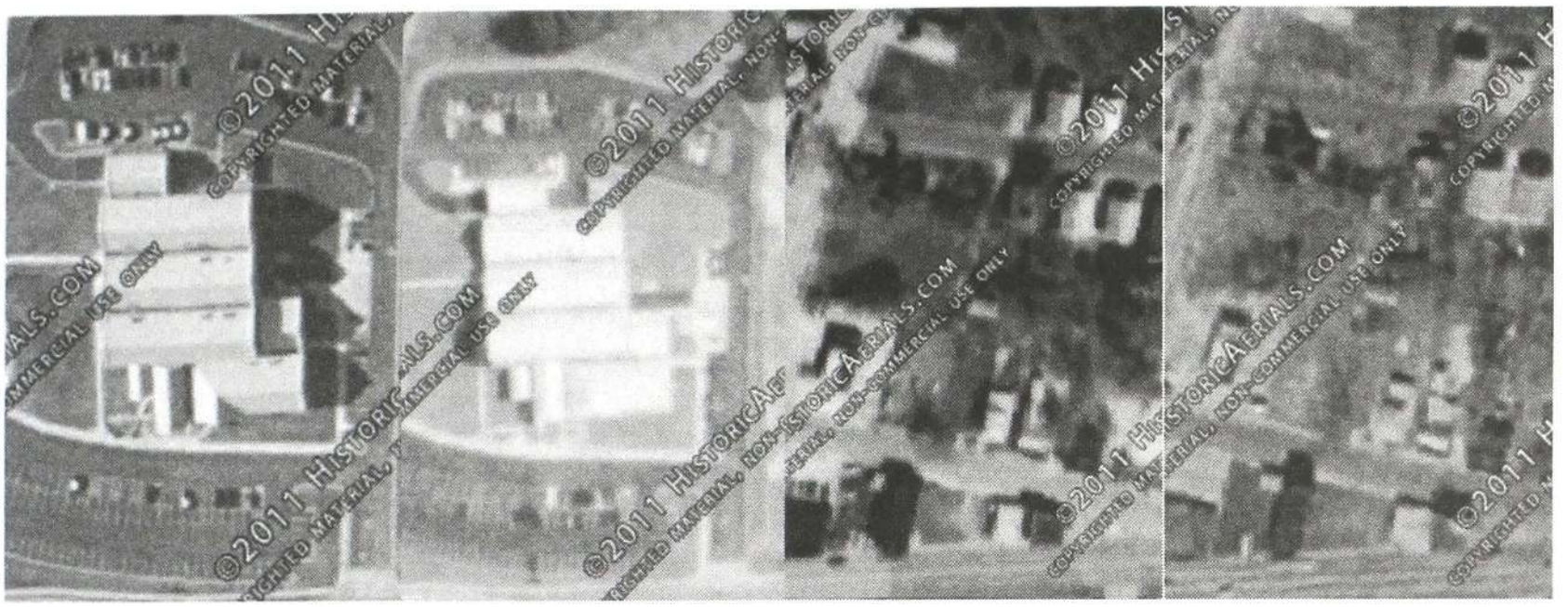


Figure 12.05: Multiple views of a location through Historic Aerials.

Satellite Imagery Update Notification (followyourworld.appspot.com)

It is safe to assume that Google will continue to generate new satellite views of earth as time passes. If you have a specific location of interest, you may want to be notified the moment that a new image is available for view. This Google service does exactly that. After you log into your Google account, you are allowed to select a specific location by address, GPS, or landmark. Google will now email you each time the satellite imagery for this location is updated. This could be useful for monitoring the remote location of a current investigation.

IntelTechniques Maps Search Tool (inteltechniques.com/OSINT/maps.html)

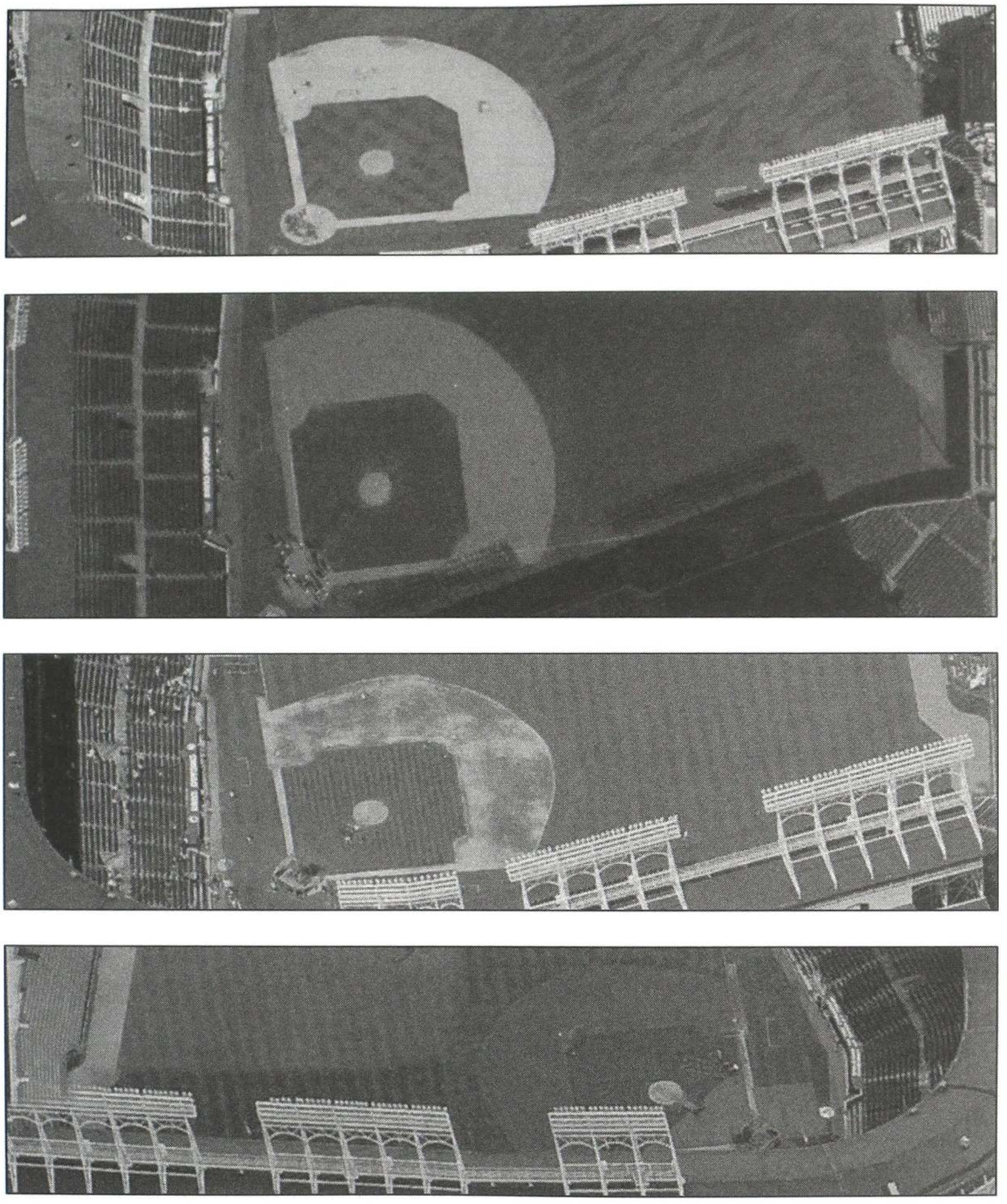
This custom search tool has two independent portions. The first option allows entry of a physical address in traditional format. This will populate the four options directly below, which execute several searches. The first presents the Google Maps API page that includes the GPS coordinates of your target. These will be used in the second portion of the page. The Zillow Homes, Rehold Homes, and Google Homes options conduct searches attempting to identify any interior images from house sale websites. These have been very valuable during briefings before a search warrant execution.

The second portion allows entry of GPS coordinates, which will populate all remaining options. Each search will open results in a new tab, or the final "Submit All" will display all satellite imagery from multiple providers within new tabs. It currently fetches images from Google Satellite

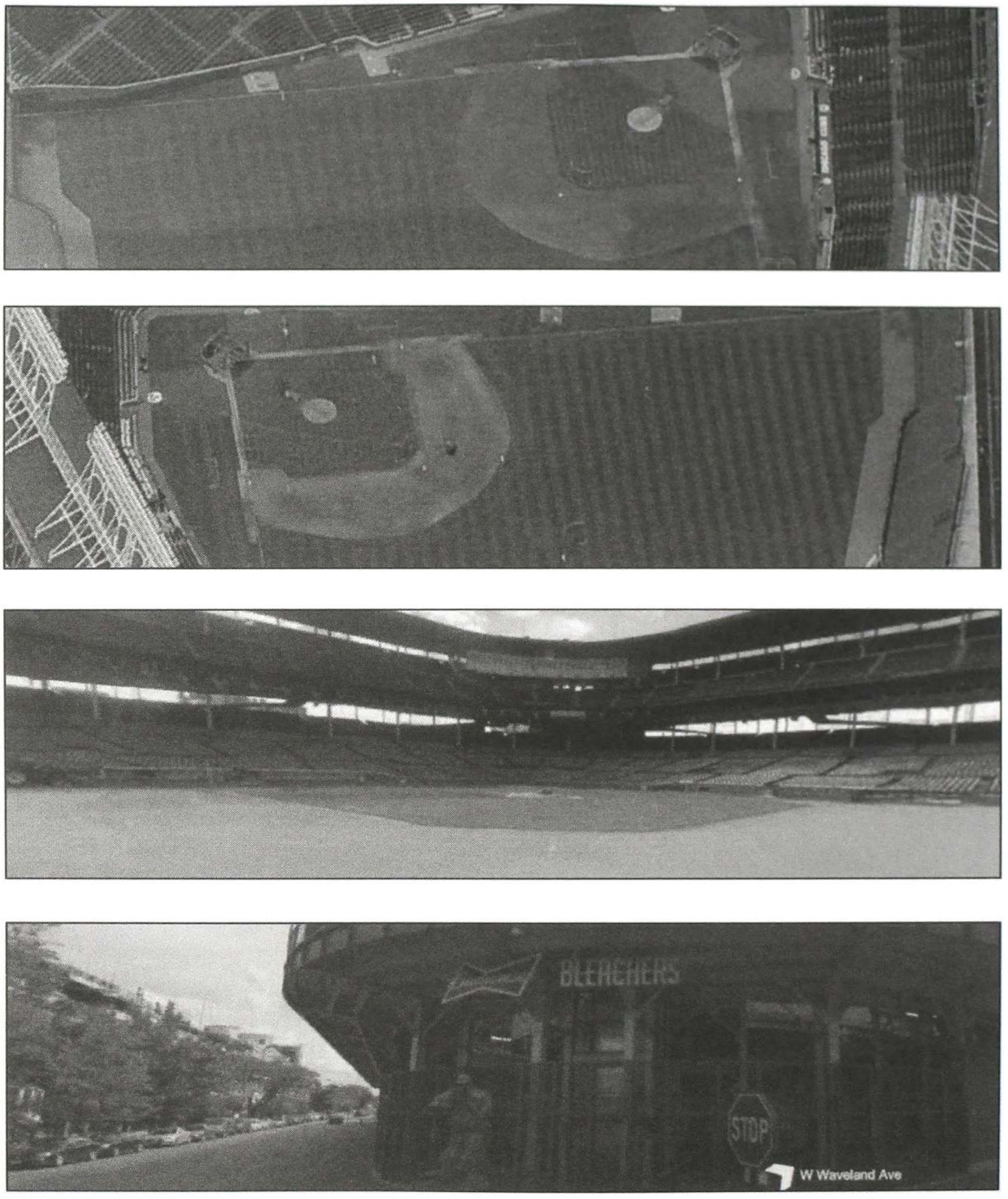
(standard and 45 degree views), Bing Satellite, Bing Bird's Eye (N, E, S, W), Google Street View, Bing Street View, Terra Server Satellite, Land Viewer Satellite, Here Satellite, Wikimapia Satellite, Zoom Earth Satellite, Yandex Satellite, Map Box Satellite, Descartes Satellite, Mapillary Street View, and Open Street Cams Street View. It also queries YouTube, Facebook, and Periscope for any posts from the target. Figure 12.06 displays the current view of the tool. Figures 12.07 through 12.20 on the following pages display the results from these providers, in order, when searching Wrigley Field in Chicago.



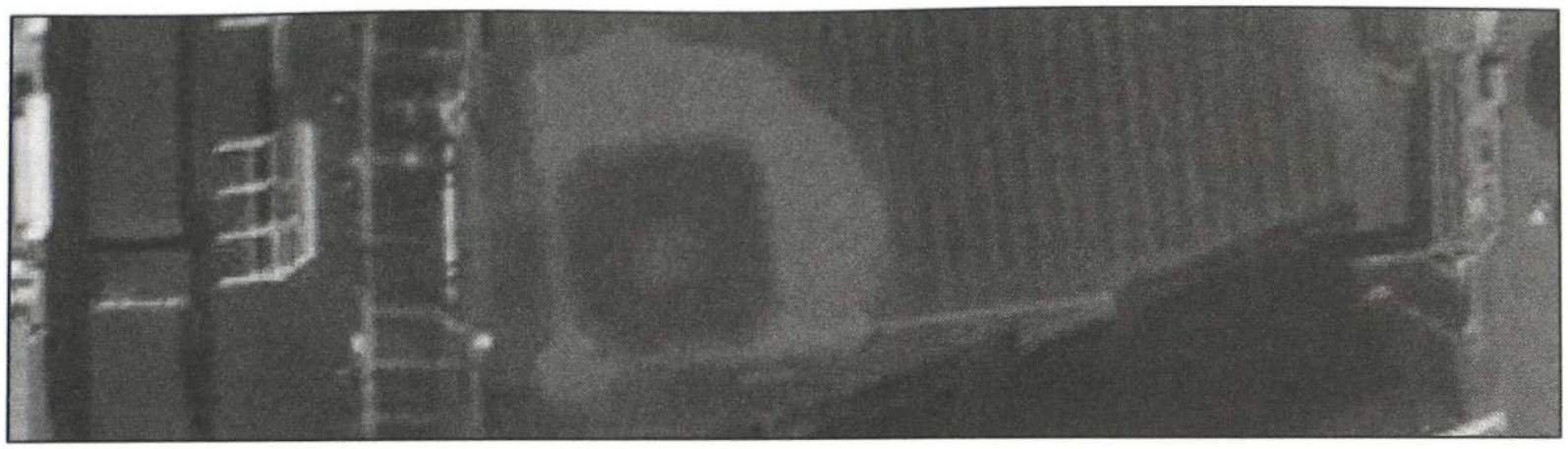
Figure 12.06: The IntelTechniques Custom Maps Search Tool.

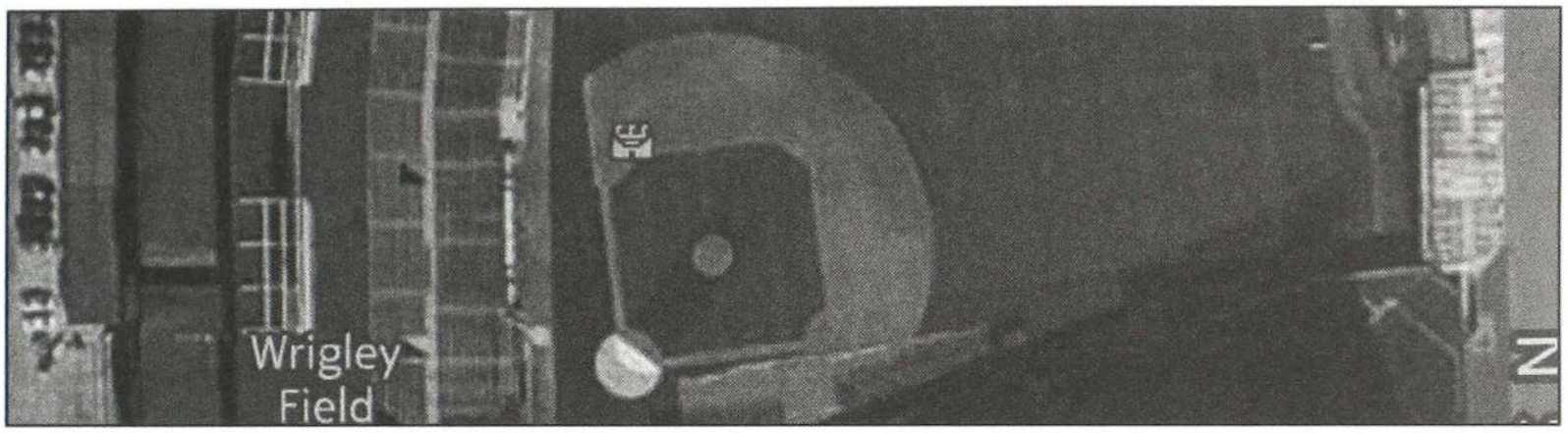


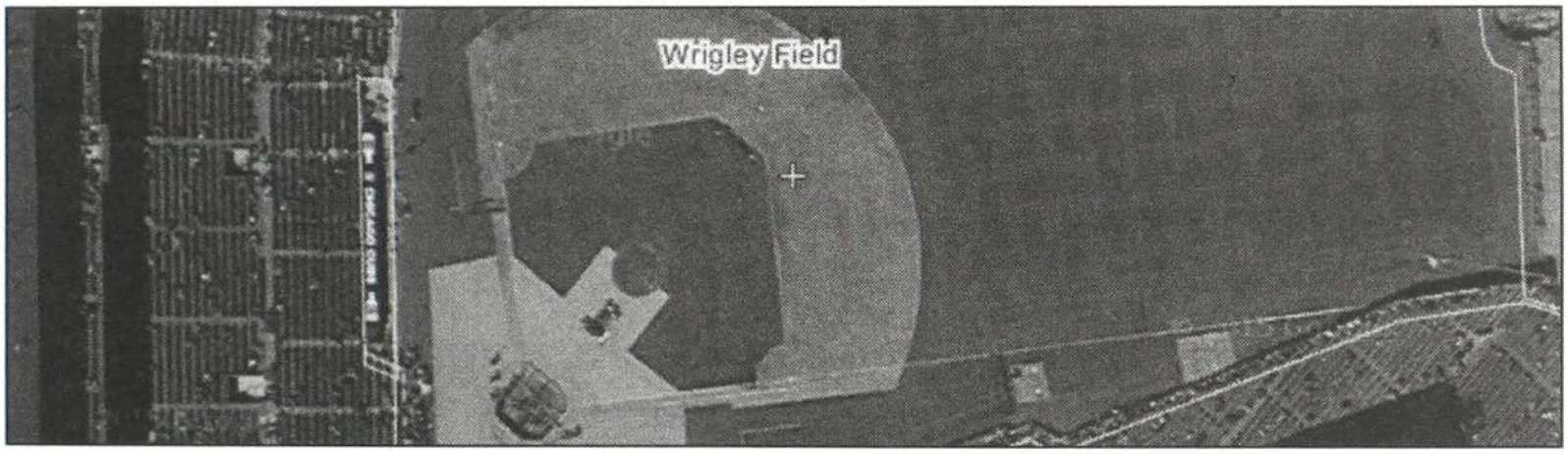
Figures 12.07 through 12.10: Satellite views from Google, Bing, Bing Bird's Eye (north), and Bing Bird's Eye (east).

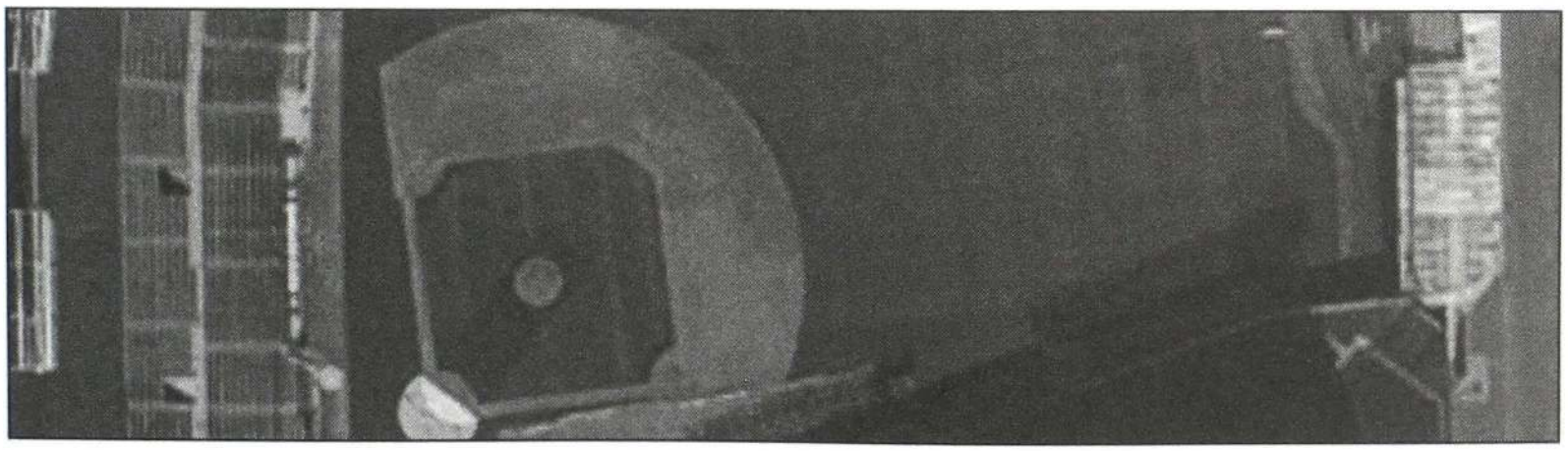


Figures 12.11 through 12.14: Satellite views from Bing Bird's Eye (south), Bing Bird's Eye (west), Google Street View, and Bing Street View.

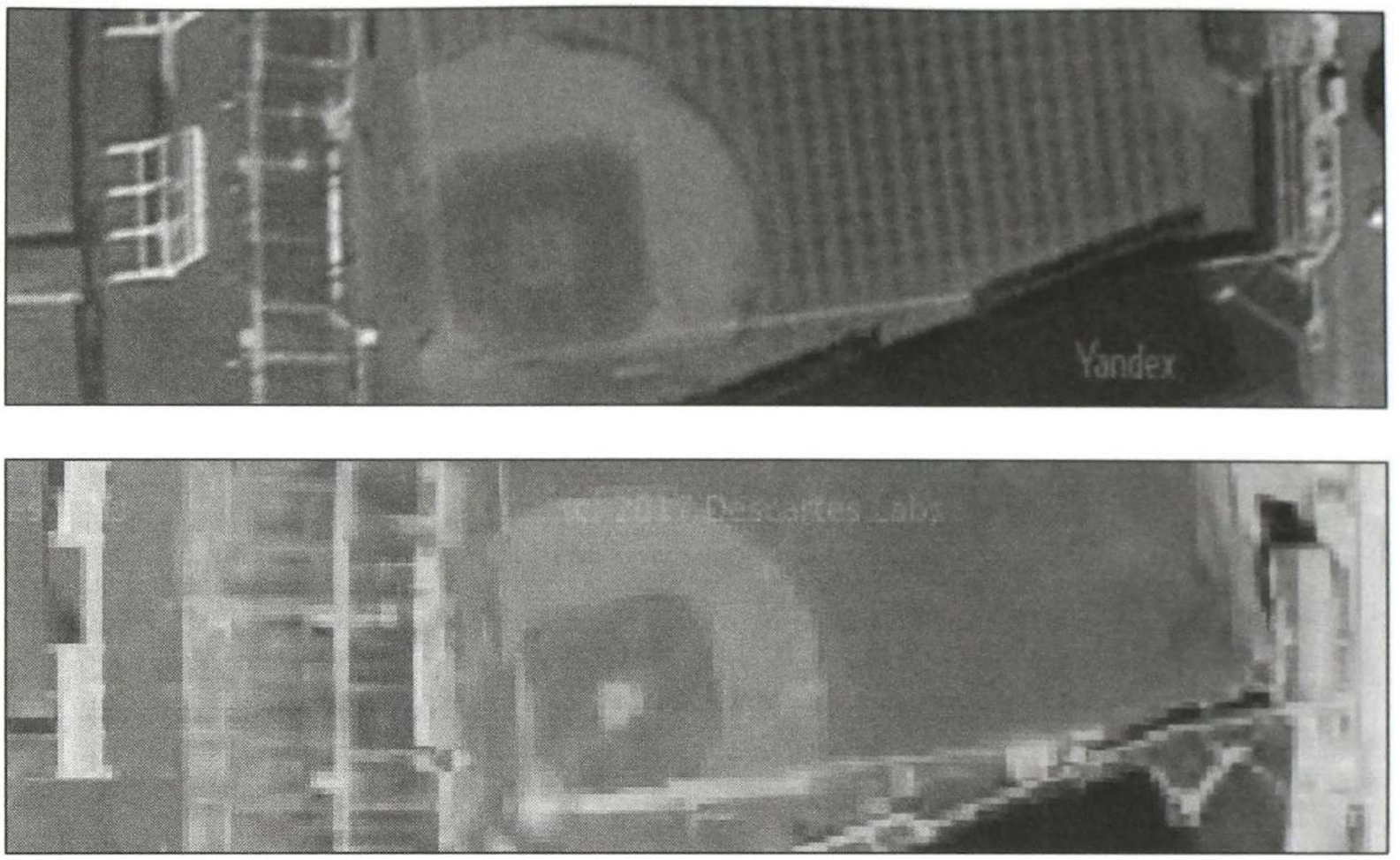








Figures 12.15 through 12.18: Satellite views from Terra Server, Here, WikiMapia and ArcGIS.



Figures 12.19 through 12.20: Unique satellite views from Yandex and Descartes.

I use this tool during practically every investigation. Whether I have a residential address of a target or a business address of the scene of an incident, these satellite views have relevance. Having fourteen unique views of an individual location would have seemed unimaginable in decades past. Today, we take it for granted. I believe that capturing these images while present is important. These services disappear for unknown reasons or images are overwritten with new content. Archiving these views during your research preserves the evidence forever.

Classic Maps (gokml.net/maps#)

In previous editions of this book, I explained the ways that users could access the original Google Maps, often referred to as the Classic Maps. There were a handful of URLs that still presented Google's classic view with additional tools. Google has officially disabled all of these hidden sites, but there is still one option left. Gokml.net hosts an interactive map tool that appears very similar to the classic version of Google Maps. It uses Google's live stream of mapping data, but displays options no longer available in the current Google Maps. Most of these are related to the layout of the page and allow you to customize the view. It allows you to split the view horizontally and vertically, with Street View on one side and Satellite View on the other. While this does not present any new content, it does provide additional view options. If this page were to ever disappear, I maintain a copy at the following address. Download this file to your computer and execute the file within your chosen browser.

https://inteltechniques.com/OSINT/iframe/classic.maps.html

Scribble Maps (scribblemaps.com)

The default view of mapping services such as Google and Bing may be enough for your situation. Occasionally, you may want to modify or customize a map for your needs. Law enforcement may want to create a map to be used in a court case; a private investigator may want to customize a map to present to a client; or a security director may want to use this service to document the inappropriate Tweets that were found during the previous instructions. Scribble Maps offers one of the easiest ways to create your own map and add any type of visual aids to the final product.

The default view of your new map at Scribble Maps will display the entire world and a menu of basic options. I close this menu by clicking the small "x" in the upper right corner. You can then manually zoom into an area of interest or type in an address in the location bar at the top of the map. This will present you with a manageable area of the map. The lower right corner will allow you to switch from a traditional map view to a satellite or hybrid view.

The menu at the top of the map will allow you to add shapes, lines, text, and images to your map. Practicing on this map can never be replaced with any instruction printed here. Mastering the basics of this application will make occasional use of it easy. Figure 12.21 displays a quick sample map that shows a title, a line, a marker, and graphics. The menu can be seen in the upper left portion. When finished, the "Menu" button will present many options to print, save, or export your map. I also highly recommend **Free Map Tools** (freemaptools.com). This service provides multiple advanced options such as mapping a radius around a point of interest.

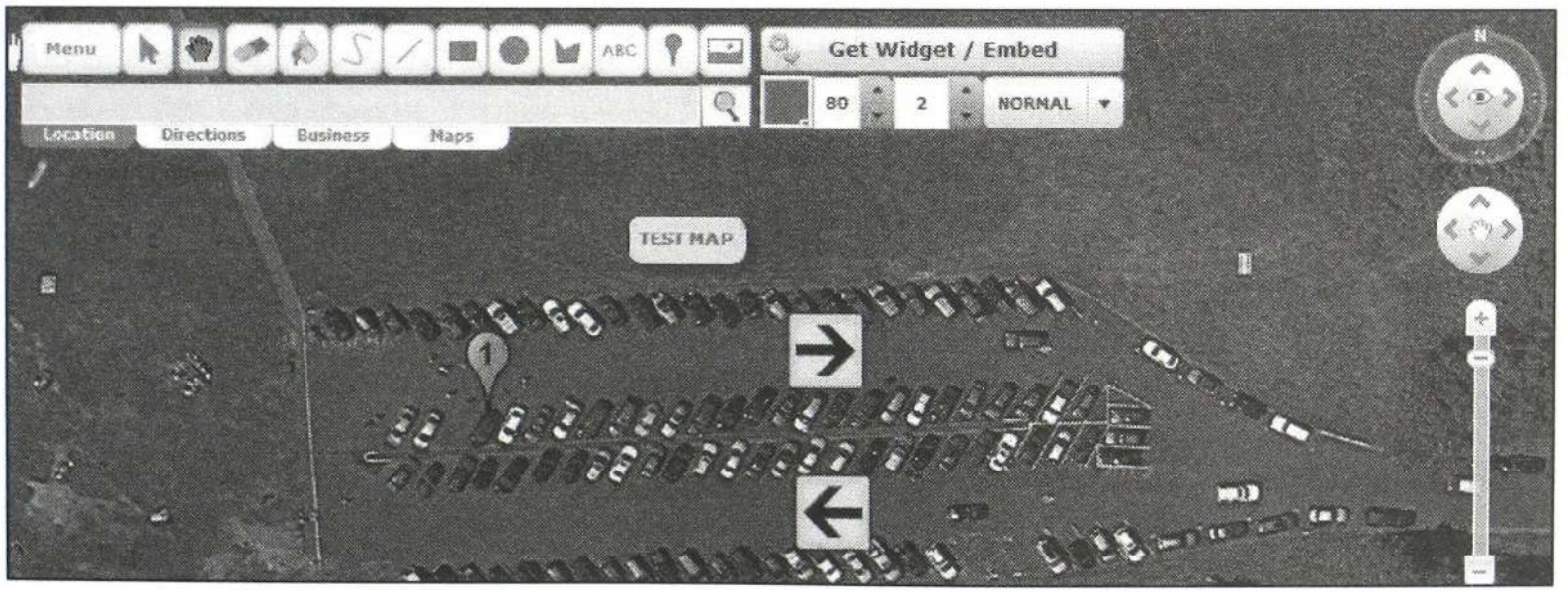


Figure 12.21: A basic custom map created with Scribble Maps.

CHAPTER THIRTEEN
DOCUMENTS

The open source intelligence discussed up to this point has focused on websites that include valuable information about a target. A category of intelligence that is often missed during OSINT research is documents. This type of data usually falls into one of three classes. The first is documents that include information about the target within the contents of the file. These can include online PDF files that the target may not know exist. The second class is documents that were actually created by the target. These files can make their way into public view unintentionally. Finally, the third class includes the metadata stored within a document that can include vital information about the true source of the document. The following techniques explain manual searching and retrieving of documents. Later in this book, automated software solutions will be detailed.

Google Searching (google.com)

A very basic way of locating documents that are publicly available on a specific website, or related to a specific topic, is to use Google. The "filetype" search operator explained in Chapter Three can be used for this task. An example of a search query for all Microsoft Word documents stored on the domain of inteltechniques.com would be the following.

site:inteltechniques.com filetype:doc or site:inteltechniques.com filetype:docx

If you wanted to locate all documents that reference a specific topic, you can use the filetype operator without a specific website listed. An example of a search query for all Excel spreadsheets that contain the acronym OSINT would be the following.

filetype:xls "OSINT"

This search yielded 82 results for Excel documents. If you wanted to search for a specific person's name within any spreadsheets, such as John Doe, you would type the following query.

filetype:xls "John Doe"

The following table includes the most common document file types and the associated file extensions. As explained in Chapter Three, both Google and Bing are capable of searching any file type regardless of the file association. Please note that this is a partial list, and I identify new possibilities constantly.

Microsoft Word	DOC, DOCX
Microsoft Excel	XLS, XLSX, CSV
Microsoft PowerPoint	PPT, PPTX
Adobe Acrobat	PDF
Text File	TXT, RTF
Open Office	ODT, ODS, ODG, ODP
Word Perfect	WPD

If you wanted to search all of these file types at once, the following string in Google or Bing would find most documents on the topic of OSINT. You could change that term to anything else of interest.

OSINT filetype:pdf OR filetype:doc OR filetype:xls OR filetype:xlsx OR filetype:docx OR filetype:ppt OR filetype:pptx OR filetype:wpd OR filetype:txt

This query basically tells the search engine to look for any reference to the term OSINT inside of a PDF file, Microsoft Word file, et cetera, and display all of the results. The Google Custom Search Engine described in Chapter Three is a great resource for this exact type of search. However, I highly recommend having an understanding of the manual search process. It will give you much more control than any automated solution. The first three editions of this book contained several third-party document search services. Most of them either disappeared or now rely solely on a Google custom engine. Therefore, I no longer recommend any of them. They simply cannot compete with a properly structured document search on Google or Bing.

Google Docs (docs.google.com)

The idea of storing user created documents on the internet is gaining a lot of popularity. Keeping these files "in the cloud" eliminates the need for personal storage on a device such as a CD or flash drive. In addition, storing files on the internet allows the author to access and edit them from any computer with an internet connection. A common use of these document-hosting sites is to store them only during the editing phase. Once the document is finished and no longer needed, the user may forget to remove it from public view. Google is one of the most popular document storage websites. It allows users to embed the stored documents into their own websites if desired. Searching the site is relatively easy.

Many Google Mail (Gmail) users take advantage of Google's free service for document storage called Google Docs or Google Drive. When a document is created, it is private by default and not visible to the public. However, when people want to share documents with friends or coworkers, the sharing properties must be changed. While it is possible to privately share files with individual Google users, many people find it easier to make the documents public. Most of these users probably assume that the files will not be seen by anyone other than the intended recipients. After all, who would go out searching for other people's documents? We will.

The Google Docs and Google Drive websites do not offer the option to search these public files, but you can do this using Google search. Now that Google Docs allows search engines to index most of the public files, you should be able to find them with some specific search methods. The following search examples will explain a few of the options that would be conducted on google.com. The exact search is listed with the expected result. These should be used as a starting point for the many possibilities of document searching.

site:docs.google.com "resume" - 22,700 online resumes
site:docs.google.com "resume" "Williams" - 2,491 resumes with the name Williams
site:docs.google.com "Corey Trager" - 1 document (resume) belonging to the target
site:docs.google.com 865-274-2074 - 1 document containing the target number

Google categorizes the documents that are created by the user. The examples below identify searches that would display documents by type.

site:docs.google.com/presentation/d – 865,6000 PowerPoint presentations
site:docs.google.com/drawings/d – 68,600 Google flowchart drawings
site:docs.google.com/file/d – 6,945,000 images, videos, PDF files, and documents
site:docs.google.com/folder/d – 94,000 collections of files inside folders
site:docs.google.com/open – 1,400,000 external documents, folders, and files

In 2013, Google began placing some user generated documents on the "drive.google.com" domain. Therefore, any search that you conduct with the method described previously should be repeated with "drive" in place of "docs". The previous search for the telephone number would be the following.

site:drive.google.com 865-274-2074

Microsoft OneDrive (onedrive.live.com)

Similar to Google Drive, Microsoft's OneDrive offers that ability to store and share documents. The service is not as popular as Google Drive. However, there are thousands of publicly visible documents waiting to be found. The shared files are stored on the onedrive.live.com domain. A query for resumes would be as follows. This search could be conducted on Google or Bing. The result on Google was 3,550 resume files with personal information.

site:onedrive.live.com "resume"

Amazon Web Services (amazonaws.com)

Amazon Web Services is a large collection of servers that supply storage and internet application hosting in "the cloud". Instead of purchasing expensive hardware, many companies and individuals rent space on these servers. There are numerous documents available for download

from these servers when searched appropriately. I cannot overstate the value of searching Amazon's servers. This is where most of the voter data that was heavily discussed during the 2016 election originated. I have personally located extremely sensitive documents from this source on numerous occasions. The following structure will identify files indexed on google.com.

site:amazonaws.com

The following search examples will explain a few of the options. The exact search is listed with the expected result. These should be used as a starting point for the many possibilities of document searching.

site:amazonaws.com ext:xls "password" – 1,001 Excel spreadsheets
site:amazonaws.com (504) 390-6582 – 1 PDF file with target number
site:amazonaws.com "lionheart201" – 1 PDF file with user name reference

Another option is the Amazon Cloudfront servers. CloudFront is a content delivery network (CDN) offered by Amazon Web Services. Content delivery networks provide a globally-distributed network of proxy servers which cache content, such as web videos or other bulky media. These are provided more locally to consumers, thus improving access speed for downloading the content. We can apply the same previous search techniques on this domain. The following search on Google yielded 57 results of pages on various Cloudfront servers containing the acronym "OSINT".

site:cloudfront.net OSINT

Presentation Repositories

With unprecedented online storage space at all of our fingertips, many people choose to store PowerPoint and other types of presentations in the cloud. Several free services have appeared to fill this demand. Of those, the following have the majority of publicly available documents.

Slide Share (slideshare.net)
ISSUU (issuu.com)
Prezi (prezi.com)

Slide Share and ISSUU allow native searching within their websites. However, Prezi does not have this option. For all three, I recommend a custom Google search with the site operator. If I want to locate presentations including the term OSINT from Slide Share, I would use the following query.

site:slideshare.net "OSINT"

Scribd (scribd.com)

Scribd was a leading cloud storage document service for several years. Since 2014, it has shifted its focus toward ebook sales. However, the plethora of stored documents is still accessible. This can be valuable for historic content posted, and likely forgotten, by the target. A search field is at the top of every page on the site within their collapsible menu. Searching for your target name should produce any public books stored through this service that includes the target name on any page of the publication. Clicking "Documents" in the menu will present more relevant information. Most of these documents are intentionally stored on the site and any groundbreaking evidence of criminal activity will not be included. Instead, the primary use of the site for OSINT investigations is the large number of documents related to businesses. Entering any large corporation name should display several pages of viewable documents related to the company. Often, these include documents that the company's security personnel would not authorize to be online. Searching for "FOUO", an acronym for "for official use only", produced hundreds of results. While none of these appeared to be officially classified, they were not intended to be posted to a public website. If you are presented with an unmanageable amount of results, the filter options appear directly above the first document result. These will allow you to search by language, size, file type, and date uploaded.

Identifying the user that uploaded a document is as easy as locating the document. In the upper-center of any page containing a document, there is an area that will identify the subject that uploaded the file. This also acts as a link to this user's profile on the website. The profile will display any information that the user supplied as well as a feed of recent activity of that user on the site. This can help identify other documents uploaded by a specific user.

IntelTechniques Documents Search Tool (https://inteltechniques.com/OSINT/docs.html)

If these operators seem overwhelming, consider the Google Custom Search Engines (CSE) that were explained in Chapter Three. They apply most of the methods discussed here. These engines will present a simple search field ready for any keyword desired. I have found an all-in-one Google CSE to be a bit unreliable and sporadic with results. Therefore, I created my own option, which can be seen in Figure 13.01. This search tool has three sections, all of which are explained below.

The "Documents by Service" column allows entry of any terms, and the first option will populate the remaining search fields. You can then choose to execute your search for data on Google Docs, Google Drive, Microsoft Drive, Amazon AWS, Cloudfront, SlideShare, Prezi, ISSUU, Scribd, and PDF Drive. The "Submit All" option will execute each search in its own tab.

The "Documents by File Type" column allows entry of any terms, and the first option will populate the remaining search fields. You can then choose to execute your search for data with specific file extensions, including PDF, DOC, DOCX, XLS, XLSX, CSV, PPT, PPTX, KEYNOTE, TXT, RTF, XML, ODT, ODS, ODP, ODG, ZIP, RAR, 7Z, JPG, JPEG, PNG, MPG, MP4, MP3, and WAV. The "Submit All" option will execute each search in its own tab.

The final two search options are Google custom search engines (CSE). The results of the Documents by Storage Service CSE will include content obtained from Google Drive, SlideShare, AmazonAWS, ISSUU, Scribd, DocStoc, and Prezi. The Search by Filetype CSE will filter results by the most common file types listed previously. Any time that I need to search for online documents, this is my first stop. I have found the specialized searching available here to rival any aggregated engines or custom search engines (CSE).

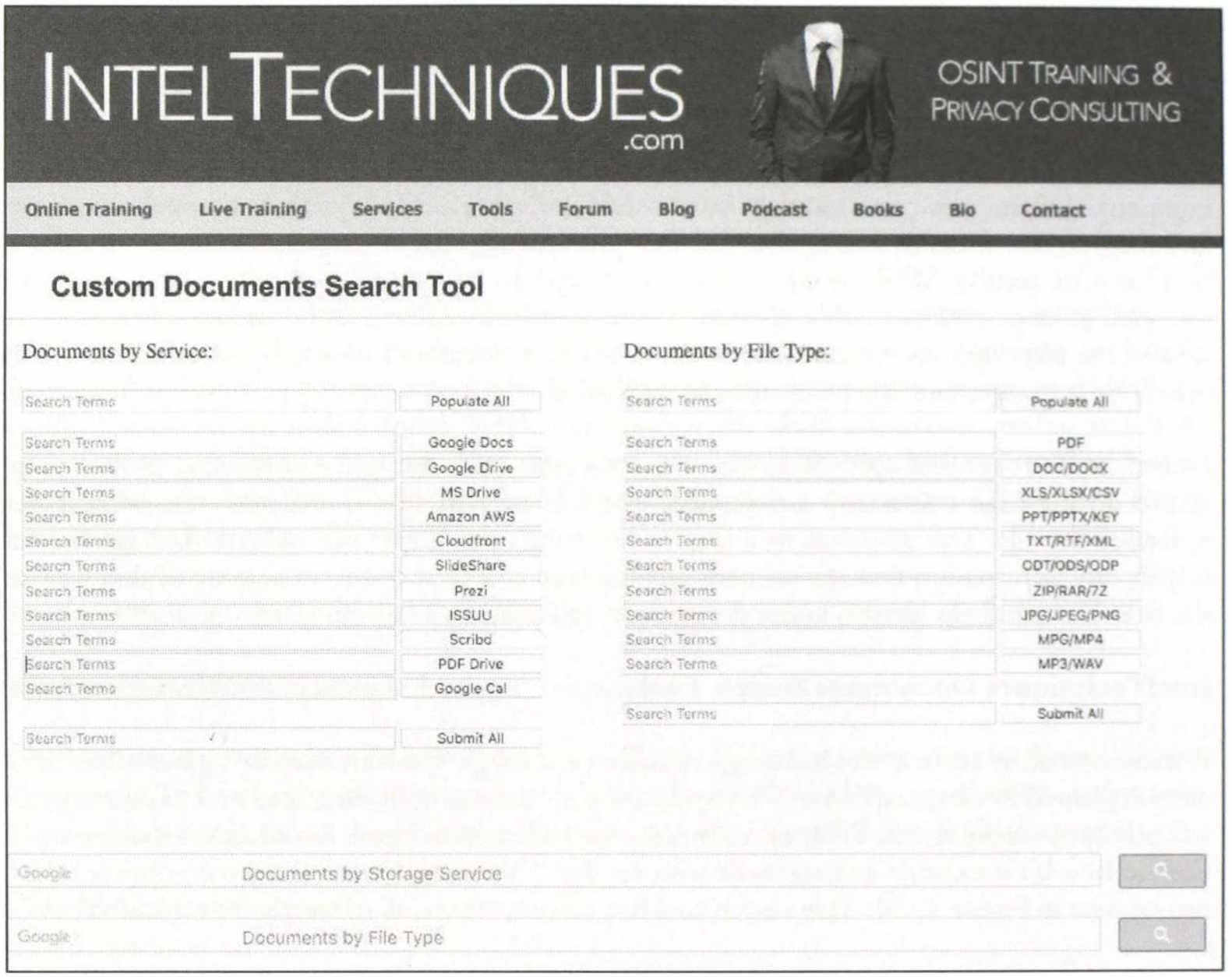


Figure 13.01: The IntelTechniques Documents Search Tool.

WikiLeaks (search.wikileaks.org)

Some websites are created for the sole purpose of leaking sensitive and classified documents to the public. Wikileaks is such a site. When an Army soldier named Bradley Manning was arrested in 2010 for uploading classified government information to the site, Wikileaks became a household name. People then began to flock to the site to catch a glimpse of these controversial documents and videos. The official Wikileaks site finally provides a working search option. It will

allow you to enter any search terms and will provide results of any leaked documents that contain these terms. Both the government and the private sector should be familiar with this site and the information that is identified with their agency.

Cryptome (cryptome.org)

Another site that strives to release sensitive and classified information to the public is Cryptome. Most of the information is related to freedom of speech, cryptography, spying, and surveillance. Much of the content could be considered conspiracy theories, but several official documents get released daily. Cryptome does not provide a search for their site and there are no third-party providers that cater to this service. Therefore, we must rely on Google or Bing to find the documents. A structured query such as the following should function well. This technique using the search terms of "bradley manning" linked to 77 documents surrounding his investigation.

site:cryptome.org "NAME OR TOPIC"

Metadata Viewers

When an original document is found online, it is obviously important to analyze the visible content of the file. This includes the file name, written text, and an original location of the document. Digging deeper will expose more information. There is data embedded inside the document that cannot be seen by simply looking at the content of the file. This data is called metadata and can be very valuable to any type of investigation. This data can often include the computer name the document was created on, the user name of the computer or the network, the software version used, and information about the network to which the computer is connected. The best way to view all this information is to use a software solution which will be discussed later in the book. It is also possible to view this "hidden" information online through a web browser.

Several online sites will allow you to upload documents for analysis. To do this, click the "browse" button on the pages detailed below. This will enable a file explorer that will allow you to select the document that you want analyzed. The result often identifies a created and modified date, the original title, three applications used to create the document, and a user name. A further search of this user name through the previously discussed techniques could produce a wealth of information about the author of the document. The following websites allow you to upload a locally stored document or submit a URL of a file for analysis. Please use caution with this technique. If the document is already posted online, there is very little risk of allowing a URL analysis. However, a locally stored file that has never been on the internet may require a second thought. If the content is sensitive, you may not want to upload to any service. If the file contains classified information, you could be jeopardizing your clearance. In these situations, use the method discussed in Chapter Nineteen. If this is not a concern the following work well.

Extract Metadata (extractmetadata.com)
Jeffrey's Exif Viewer (regex.info/exif.cgi)
Metashield Analyzer (metashieldanalyzer.elevenpaths.com)

Real World Application: Dennis Lynn Rader, also known as the BTK killer, sent a floppy disk to the Wichita Police Department containing a Microsoft Word document in reference to his killings. The police examined the metadata of this document and determined that it was made by a subject named "Dennis". Links to a Lutheran church were also located within this data. Conducting OSINT searches on these two pieces of information helped to identify the suspect and make an arrest.

Free OCR (free-ocr.com)

You may occasionally locate a PDF file that has not been indexed for the text content. These types of PDF files will not allow you to copy and paste any of the text. This could be due to poor scanning techniques or to purposely prohibit outside use of the content. You may desire to capture this text for a summary report. These files can be uploaded to Free OCR and converted to text documents. OCR is an acronym for optical character recognition. Basically, a computer "reads" the document and determines what the text is inside the content. The result is a new document with copy and paste capability.

Rental Vehicle Records

The details of rental vehicles are not technically documents, but the data seemed to fit this category the best. The following options have been controversially received during training and may not be appropriate for everyone. I present these methods to you as theories, and you should evaluate if the techniques are suitable for your research. Several vehicle rental companies offer an option to access your receipts online. This is probably designed for customers that leave a vehicle at the business after hours and later need a receipt. While the processes to retrieve these documents are designed to only obtain your own records, it is very easy to view others.

Enterprise (enterprise.com)

At the bottom of every Enterprise web page is an option to "Print your receipt". Clicking this will present a form that must be completed before display of any details. Enterprise will need the target's country, driver's license number, and last name. Providing this information will display the user's entire rental history for the past six months. Testing with my own data provided two years' worth of results. Each document will link to the entire receipt from that rental. These receipts include the start date and time, end date and time, vehicle make and model, location picked up, total mileage, lease name, and form of payment. This information could be very beneficial to any drug case or private investigation.

Hertz (hertz.com)

Similar to Enterprise, Hertz has a link at the bottom of every page titled "Find a receipt". You can search by driver's license number or credit card number and will need a matching last name. The receipt will be very similar to the example in the Enterprise demonstration.

Alamo (Alamo.com)

Alamo also titles their receipt retrieval link "Find a Receipt" and it is located in the lower right portion of every page. The process is identical to the previous two examples. The only difference is that you must choose a date range. I usually select a start date of one year prior to the current date and the end date of the current date.

Paste Sites

Paste Sites are not technically documents. They are websites that allow users to upload text for public viewing. These were originally designed for software programmers that needed a place to store large amounts of text. A link would be created to the text and the user could share the link with other programmers to review the code. This is still a common practice, but other users have found ways to abuse this technology. Many hacking groups will use this area of the internet to store compromised account information, user passwords, credit card numbers, and other sensitive content. There are dozens of sites that cater to this need, and very few of them have a search feature.

Pastebin (pastebin.com)

Pastebin is the most popular paste site in the United States. Criminal hacker groups often use this site to release illegally obtained data to the public. A recent release included the home addresses and personal information of many police officers near Ferguson, Missouri. This is one of the sites that will allow for a search from within the site. This function performs a search through Google in the same way we could with the "site" operator. Typing in a target name, email address, or business name may reveal private information not intended for the public. For law enforcement, typing in the last four digits of a stolen credit card number may identify a link to the thief. If successful, the target is most likely outside of the country. Regardless, this is a valuable piece to the case and an impressive explanation to the victim. Unfortunately, most of the users leave a default user name of "Guest".

IntelTechniques Paste Search Tool (inteltechniques.com/OSINT/pastebins.html)

There are dozens of online paste sites and more are added monthly. Searching all of them can be overwhelming. A custom search website that queries all known paste sites is located at the above address. At the time of this writing, it searched 57 websites commonly used to host criminal information. A complete list of sites can be found on the tool.

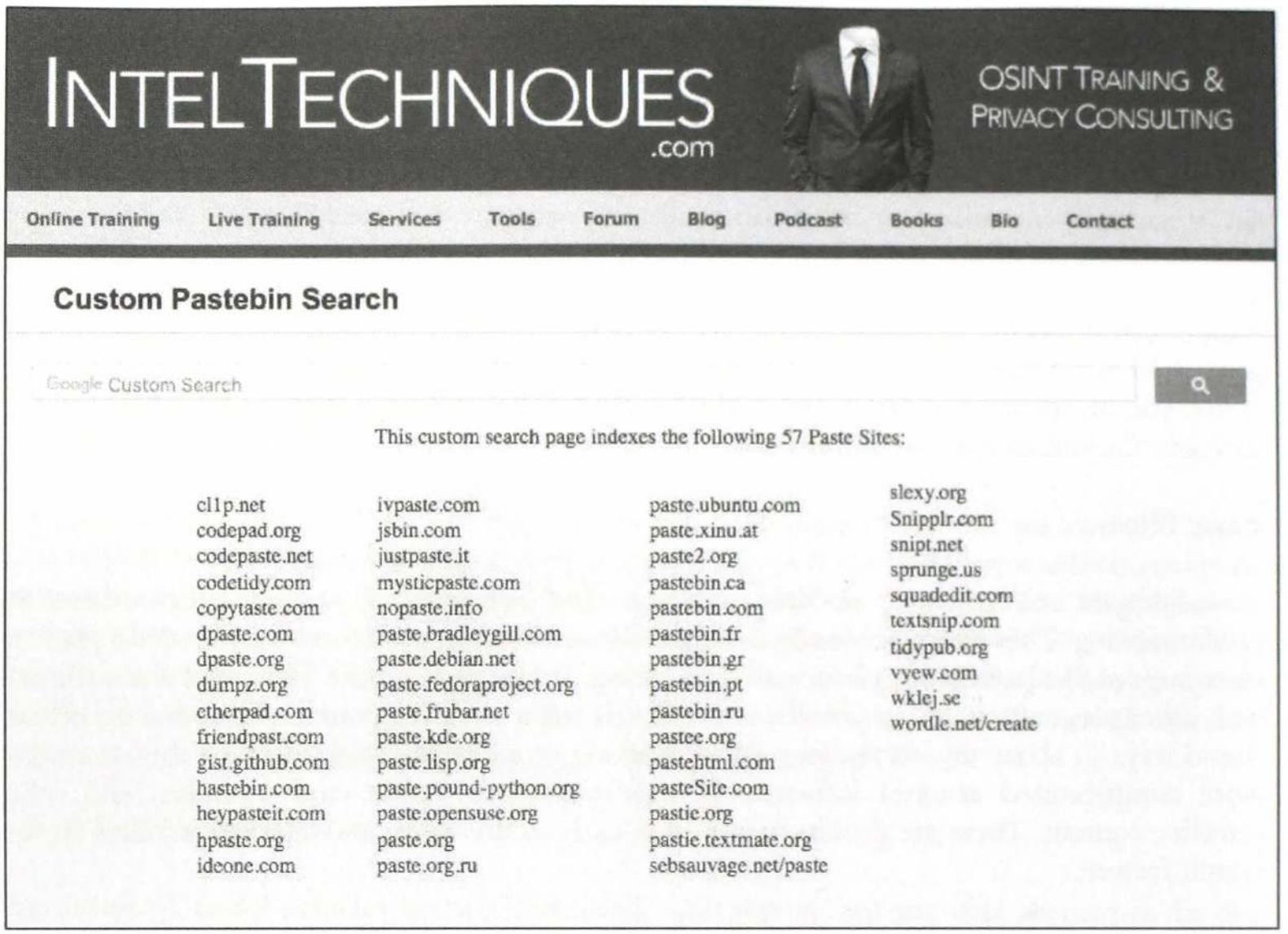


Figure 13.02: The IntelTechniques Paste Search Tool.

CHAPTER FOURTEEN
PHOTOGRAPHS

Thanks to cameras on every data cellular phone, digital photograph uploads are extremely common among social network users. These images can create a whole new element to the art of open source intelligence analysis. This chapter will identify various photo sharing websites as well as specific search techniques. Later, photo metadata will be explained that can uncover a new level of information including the location where the picture was taken, the make, model and serial number of the camera, original uncropped views of the photos, and even a collection of other photos online taken with the same camera. After reading this information, you should question if your online photos should stay online.

Google Images (images.google.com)

During my live training sessions, I always encourage attendees to avoid individual searches on various photo sharing websites such as Flickr or Picasa. This is because most of these searchable sites have already been indexed by Google and other search engines. Conducting a search for "Oakland Protest" on Flickr will only identify images on that specific service that match. However, conducting the same search on Google Images will identify photos that match the terms on Flickr and hundreds of additional services. Similar to Google's standard search results, you can use the Search Tools to filter by date. Additionally, you can further isolate target images by size, color, and type, such as photographs versus line drawings. I no longer conduct manual searches across the numerous photo sharing sites. Instead, I start with Google Images.

Bing Images (bing.com/images)

Similar to Google, Bing also offers an image search. While it is not as beneficial as the Google option, it should never be overlooked. On several occasions, I have located valuable pictorial evidence on Bing that was missing from Google results. The function is identical, and you can filter search results by date, size, color, type, and license type. When searching for relevant data about a target, I try to avoid any filters unless absolutely necessary. In general, we always want more data, not less. The search techniques explained in Chapter Three all apply to queries on Google Images and Bing Images.

Reverse Image Searches

Advancements in computer processing power and image analysis software have made reverse image searching possible on several sites. While a standard search online involves entering text into a search engine for related results, a reverse image search provides an image to a search engine for analysis. The results will vary depending on the site used. Some will identify identical

images that appear on other websites. This can be used to identify other websites on which the target used the same image. If you have a photo of a target on a social network, a reverse analysis of that photo may provide other websites on which the target used the same image on. These may be results that were not identified through a standard search engine. Occasionally, a target may create a website as an alias, but use an actual photo of himself. Unless you knew the alias name, you would never find the site. Searching for the site by the image may be the only way to locate the profile of the alias. Some reverse image sites go further and try to identify other photos of the target that are similar enough to be matched. Some will even try to determine the sex and age of the subject in the photo based on the analysis of the image. This type of analysis was once limited to expensive private solutions. Now, these services are free to the public.

Google Reverse Image Search (images.google.com)

One of the more powerful reverse image search services is through Google. Rolled out in 2011, this service is often overlooked. On any Google Images page, there is a search field. Inside this field on the far right is a light grey camera icon that appears slightly transparent. Figure 14.01 (first) displays this search field. Clicking on this icon will open a new search window that will allow for either an address of an online image, or an upload of an image file on your computer. In order to take advantage of the online search, you must have the exact link to the actual photo online. Locating an image within a website is not enough. You will want to see the image in a web browser by itself, and then copy the address of the image. If I want to view the image from the actual location, I must right-click on the image and select "view image" with my Firefox browser. Chrome users will see "open image in new tab" and Internet Explorer users will see "properties" which will identify the URL of the image. This link is what you want in order to conduct a reverse image analysis. If you paste this link in the Google Images reverse online search, the result will be other similar images, or exact duplicate images, on other sites. Visiting these sites provides more information on the target.

Note that adding context to the reverse-search field after submission can improve accuracy. As an example, a reverse-search of a photo from LinkedIn might produce many inaccurate results, but including the name or employer of your target will often display only applicable evidence.

Another way to use this service is to search for a target within the Google Images search page. The images in the results will present additional options when clicked. A larger version of the image will load inside a black box. The three options to the right of the image will allow you to visit the page where the image is stored, view the image in full size, or "Search by image". Clicking the "Search by image" link will present a new search results page with other images similar to the target image. These connect to different websites which may contain more intelligence about the subject.

Bing Reverse Image Match

In 2014, Bing launched its own reverse image search option titled "Image Match". This feature

can be launched from within any page on Bing Images by clicking the Image Match icon to the right of the search field. Figure 14.01 (second) displays this option. This service does not seem to be as robust as Google's. In my experience, I often receive either much fewer results, although they do match. On a few occasions, I have received matched images that Google did not locate.

TinEye (tineye.com)

TinEye is another site that will perform a reverse image analysis. These results tend to focus on exact duplicate images. The results here are usually fewer than those found with Google. Since each service often finds images the others do not, all should be searched when using this technique. Figure 14.01 (third) displays the search menu. The icon on the left prompts the user to provide a location on the hard drive for image upload while the search field will accept a URL.

Yandex Images (images.yandex.com)

Russian search site Yandex has an image search option that can conduct a reverse image search. Similar to the other methods, enter the full address of the online image of interest and search for duplicate images on additional websites. In 2015, Yandex began allowing users to upload an image from their computers. Overall, these results will be limited. However, this option is vital for any international investigations. Figure 14.01 (fourth) displays the reverse image search icon in the far-right portion.

Baidu Images (image.baidu.com)

Similar to Yandex, the Chinese search engine Baidu offers a reverse image search. Baidu currently offers no English version of their website and only presents Chinese text. Navigating to the above website offers a search box that contains a small camera icon to the right. Clicking this presents options for uploading an image (button to left) or providing the URL of an online image within the search field itself. The results will identify similar images on websites indexed by Baidu. Figure 14.01 (fifth) displays the search page only available in Chinese.

Regardless of the services that you are executing, I urge you to use caution with sensitive images. Similar to my view of analyzing online documents for metadata, I believe that submitting online photos within these engines is harmless. If the photo is already publicly online, there is very little risk exposing it a second time. My concern involves child pornography and classified photos. As a former child pornography investigator and forensic examiner, there were several times that I wanted to look for additional copies of evidence online. However, I could not. Even though no one would know, and the photos would never appear any place they should not, conducting reverse image searches of contraband is illegal. It is technically distributing child pornography (to Google). While working with a large FBI terrorism investigation, I had possession of ten photos on which I wanted to conduct a reverse image search. The photos were part of a classified case, so I could not. Overall, never submit these types of photos from your hard drive. It will always come back to haunt you.

Whenever I have any public images that call for reverse image searching, I always check all five of these services. While I rarely ever get a unique result on Baidu, it only takes a few seconds to check every time. This diligence has paid off in the past. These manual searches do not need to be as time consuming as one may think. We can automate much of this process to save time and encourage thorough investigations. First, we should take a look at direct URL submission. For the following examples, assume that your target image is the cover of this book from the web page at inteltechniques.com/intel/book1.html. The actual target image is stored online at the URL of https://inteltechniques.com/img/osint.cover.med.jpg. The following direct addresses would conduct a reverse image search at each service listed.

Google: https://www.google.com/searchbyimage?site=search&sa=X&image_url= https://inteltechniques.com/img/osint.cover.med.jpg

Bing: http://www.bing.com/images/searchbyimage?FORM=IRSBIQ&cbir=sbi&imgurl= https://inteltechniques.com/img/osint.cover.med.jpg

TinEye: http://www.tineye.com/search/?url=https://inteltechniques.com/img/osint.cover.med.jpg

Yandex: https://www.yandex.com/images/search?img_url=https://inteltechniques.com/img/osint.cover.med.jpg &rpt=imageview

Baidu: https://image.baidu.com/pcdutu?queryImageUrl=https://inteltechniques.com/img/osint.cover.med.jpg

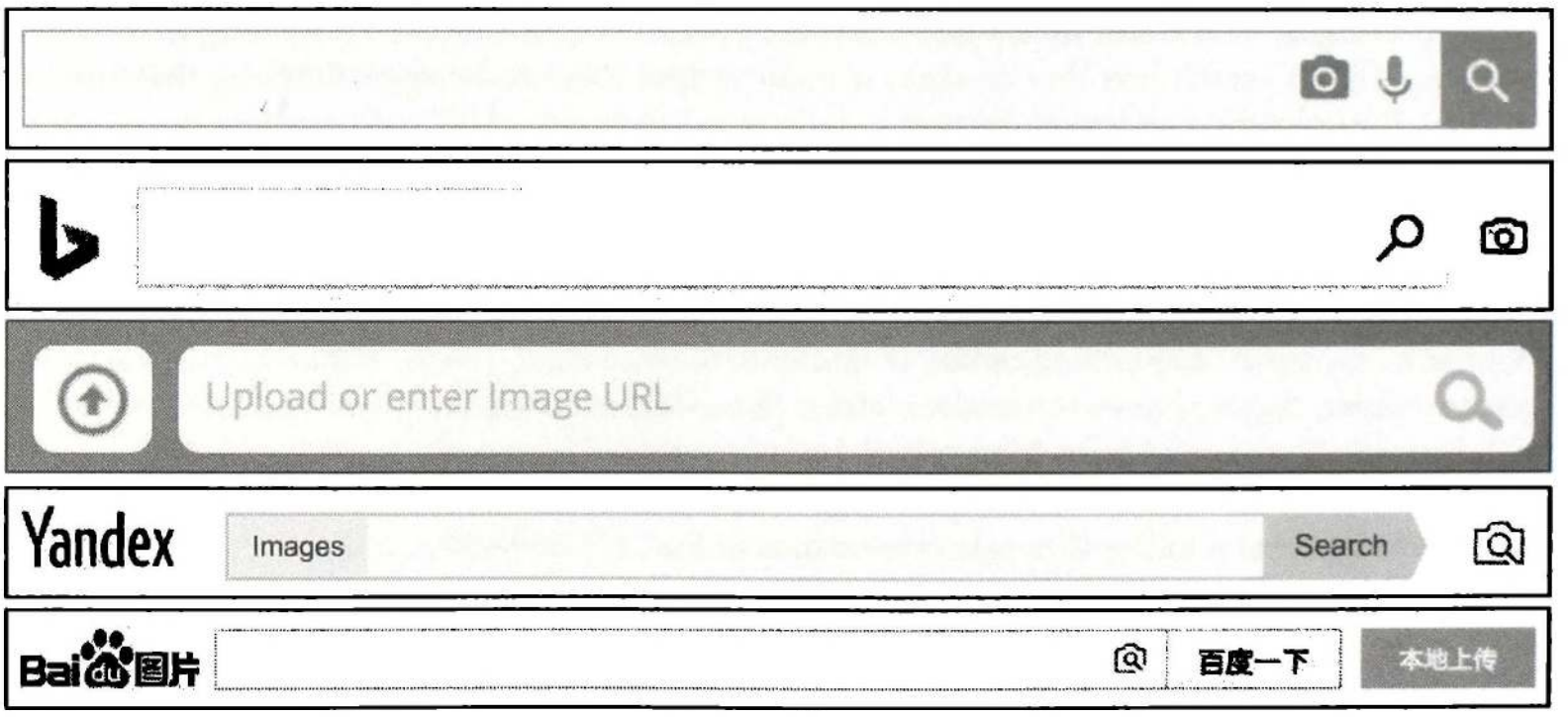


Figure 14.01: Reverse image search options from Google, Bing, TinEye, Yandex, and Baidu.

IntelTechniques Reverse Image Search Tool (inteltechniques.com/OSINT/reverse.image.html)

I do not recommend manually typing this all into a web browser. It would be more efficient to navigate to each image search site and paste the photo URL. However, I have created an online tool that automates this entire process. The first field allows input of the entire address of an online image. The Populate All button will supply this data to all fields. The next search options replicate the techniques explained here for Google, Bing, TinEye, Yandex, and Baidu. The final option on this page executes the above searches across all five networks into five separate tabs on your browser. Figure 14.02 displays the current state of this tool.

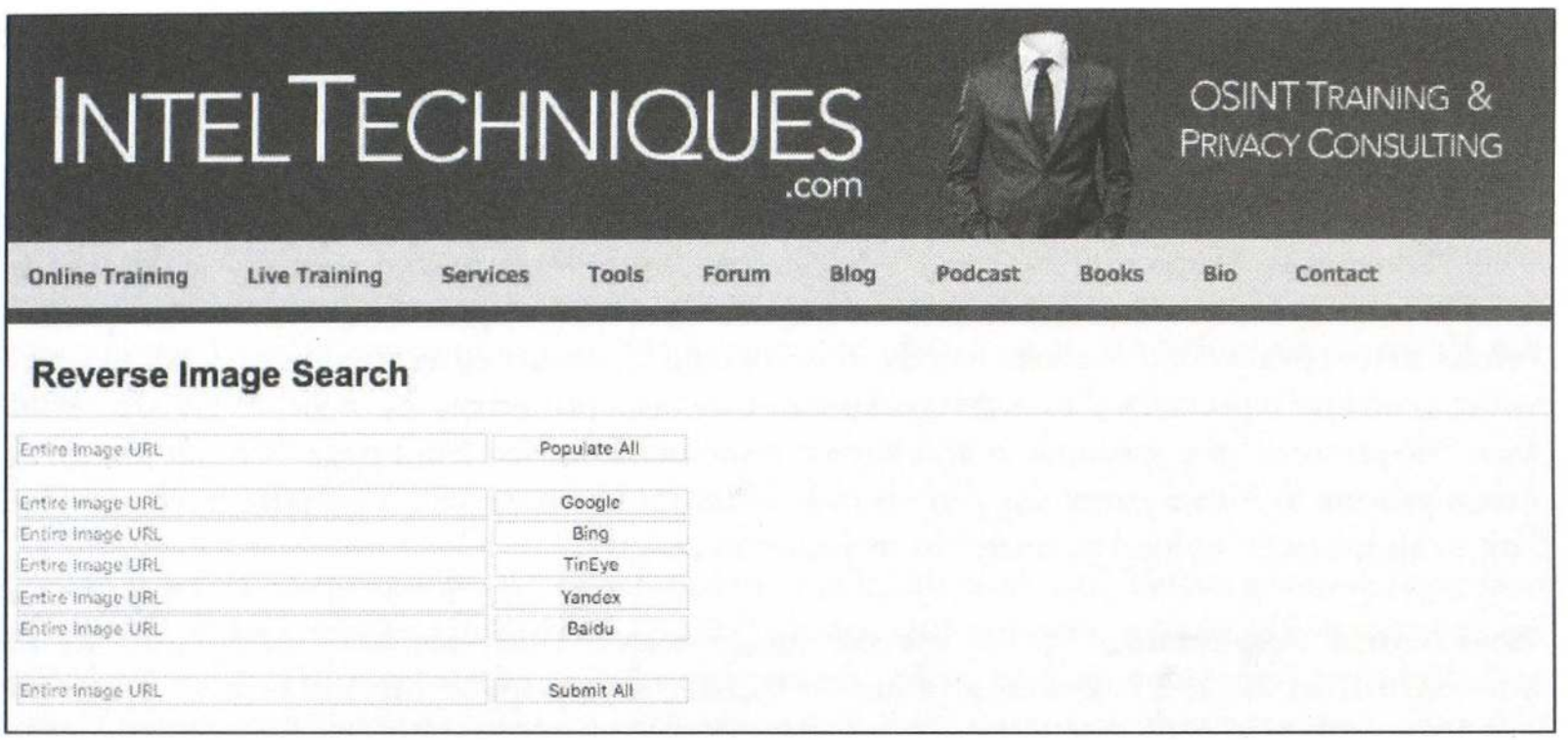


Figure 14.02: The IntelTechniques Reverse Image Search Tool.

Plag Hunter (plaghunter.com)

This German company offers a reverse image search monitoring service. It is targeted toward individuals wishing to monitor licensed images for unauthorized distribution, but investigators can take advantage of the free trial. It could be useful after identifying an online image that may be posted in the future, such as a photo of a missing child. It could also be used to monitor images of known human trafficking victims as they are posted across various prostitution forums. After creating a free account, log in and navigate to the dashboard of your portal. On the Images tab, you can provide a URL of an online image for monitoring. The page will immediately identify related images and the source of the media. The service will continue to scour the internet for images and send an email when anything new is located. At the time of this writing, the free trial was limited to five images, but the continuous scanning for these files appeared to be unlimited.

Image Raider (imageraider.com)

Image Raider is another reverse image lookup engine worth mentioning. It uses Google, Bing, and Yandex reverse image searching to provide results. However, there are some additional features which makes this tool noticeable. Image Raider lets you input up to 20 images at a time, which means you can run a multi-reverse image search by using this tool. Unlike other image search tools where you can only provide images via URL or by directly uploading from the computer, this service has multiple options for providing the source image. Among other methods, you can add a URL of a web page and it will fetch all the images from that page and use them as input for running the searches. You can also add images directly from your Flickr, DeviantArt or 500px account. Long-term use of Image Raider will cost money, but free searching from the home page should suffice for most investigations.

Karma Decay (karmadecay.com)

This service was mentioned in Chapter Seven and has a very specific specialty which can be beneficial to an internet researcher. It is a reverse image search engine that only provides positive results that appear on the website Reddit. It was originally launched as a way for users to identify when someone reposted a photo that had previously been posted on the website. The user could then "down-vote" the submission and have it removed from the front page. We can use this in investigations to locate every copy of an individual photo on Reddit. You can either provide a link to an image or upload an image from your computer.

Real World Application: These reverse image search sites can have many uses to the investigator. In 2011, I searched a photo of damage to a popular historic cemetery that was vandalized. The results included a similar photo of the suspect showing off the damage on a blog. An arrest and community service soon followed. While working with a private investigator, I was asked to locate any hotels that were using the client's hotel images on websites. A reverse image search identified dozens of companies using licensed photos without authorization. This likely led to civil litigation. In 2013, a federal agent asked me to assist with a human trafficking case. They had a woman in custody that spoke little English. She was arrested during a prostitution sting and was suspected of being a victim of trafficking. A reverse image search from one online prostitution ad located all of her other ads which identified the regional areas that she had recently been working, a cellular telephone number connected to her pimp, and approximate dates of all activity.

Pictriev (pictriev.com)

Pictriev is a service that will analyze a photo including a human face and try to locate additional images of the person. The results are best when the image is of a public figure with a large internet presence, but it will work on lesser-known subjects as well. An additional feature is a prediction of the sex of the target as well as age.

Twitter Images

For the first several years of Twitter's existence, it did not host any photos on its servers. If a user wanted to attach a photo to his or her post, a third-party photo host was required. These have always been free and plentiful. Often, a shortened link was added to the message, which forwarded to the location of the photo. Twitter now hosts photos used in Twitter posts, but third-party hosts are still widely used. The majority of the images will be hosted on Instagram, which was explained in Chapter Six and detailed further in upcoming pages. If you have already identified your target's Twitter page, you will probably have the links you need to see the photos uploaded with his or her posts.

Many Twitter messages have embedded images directly within the post. Twitter now allows you to search keywords for photo results. After you conduct any search within the native Twitter search field, your results will include a filter menu on the top. The results will only include images that have a reference to the searched keyword within the message or hashtag. You can also filter this search for people, videos, or news. Chapter Five already explained the process of obtaining copies of the images in the highest resolution possible. If you do not have an identified target yet, you will need a search engine to find the images. Standard Google and Bing searches will not always suffice.

Twicsy (twicsy.com)

Twicsy is a very user-friendly site, which boasts that it can search all Twitter photos, regardless of host. They have indexed over 6,000,000,000 photos and estimate several million new photos added every day. The top search bar is the only search option and can handle any term including topics, names, and locations. Clicking each image will direct you to another page on Twicsy that will display the full Twitter message, a larger version of the image, and additional images uploaded by the same user. To navigate to the original host of the image, in order to download the largest copy possible, click on the link within the text of the post that forwards to the photo-sharing site. As explained in Chapter Five, Twicsy can be a great repository of images that were intentionally deleted by a user. Just because a target removed posts and images from their Twitter account, it does not mean that Twicsy has erased their copies.

Twipho (twipho.net)

Twipho does not search all of the Twitter photos, but it is a decent alternative to Twicsy. Additionally, it allows for a view of a live stream of recent Twitter images being posted. Logging into a Twitter account is required for this service. The "Find photos near me" option will identify your location based on your IP address and display Twitter photos that were tagged in that general area. There is no advanced filter for this. If you are searching a populated area, this is not usually beneficial. However, searching your local rural area could provide great results.

Photo-Sharing Sites

In order to find a photo related to a target, the image must be stored on a website. The most common type of storage for online digital photos is on a photo-sharing site. These sites allow a user to upload photographs to an account or profile. These images can then be searched by anyone with an internet connection. Almost all of these hosts are free for the user and the files will remain on the site until a user removes them. There are dozens of these services, many allowing several gigabytes worth of storage. While I mentioned earlier that a Google Images or Bing Images search was most appropriate for all photo sharing hosts, Flickr deserves a mention.

Flickr (flickr.com)

Flickr, now owned by Yahoo, was one of the most popular photo-sharing sites on the internet. Many have abandoned it for Twitter and Instagram, but the mass amount of images cannot be ignored. The majority of these images are uploaded by amateur photographers and contain little intelligence to an investigator. Yet there are still many images in this "haystack" that will prove beneficial to the online researcher. The main website allows for a general search by topic, location, user name, real name, or keyword. This search term should be as specific as possible to avoid numerous results. An online user name will often take you to that user's Flickr photo album.

After you have found either an individual photo, user's photo album, or group of photos by interest, you can begin to analyze the profile data of your target. This may include a user name, camera information, and interests. Clicking through the various photos may produce user comments, responses by other users, and location data about the photo. Dissecting and documenting this data can assist with future searches. The actual image of these photos may give all of the intelligence desired, but the data does not stop there. A search on Flickr for photographs related to the Occupy Wall Street protesters returned over 157,000 results.

My Pics Map (mypicsmap.com)

Flickr is one of the few remaining web services that does not remove the metadata stored within images. The data, which will be explained in detail in a moment, can often identify the location where a photo was captured. While Flickr offers their own interactive map to display the locations of user's photos, it seldom works properly. My Pics Map is a good alternative for this type of query. The home page will allow you to enter either a Flickr user number or photoset ID number. I usually only use this to see the locations of a specific Flickr user's photos. You can zoom into each area to see exact detail of each photo and location. This can be a great resource to quickly determine the areas where a target has visited.

Flickr Map (flickr.com/map)

Flickr attempts to geo locate all of the photos that it can. It attempts to identify the location where the photo was taken. It will usually obtain this information from the Exif data, which will

be discussed in a moment. It can also tag these photos based on user provided information. Flickr provides a mapping feature that will attempt to populate a map based on your search parameters. I have had rare success with this technique. For the mapping of Flickr photos, I prefer EchoSec. This amazing free service was explained in Chapter Five.

IntelTechniques Social Images Search Tool (inteltechniques.com/intel/osint/images.html)

Similar to the search engines tool, social networks tool, and documents search tool, I also maintain a custom image search tool. This is a custom search engine (CSE) from Google and it both searches and filters results for Twitter, Facebook, Instagram, LinkedIn, Google+, MySpace, Flickr, and Picasa. Figure 14.03 displays this tool banner with a search of the term OSINT. This example identifies 45 images on Instagram with this keyword. The results will appear below the banner.

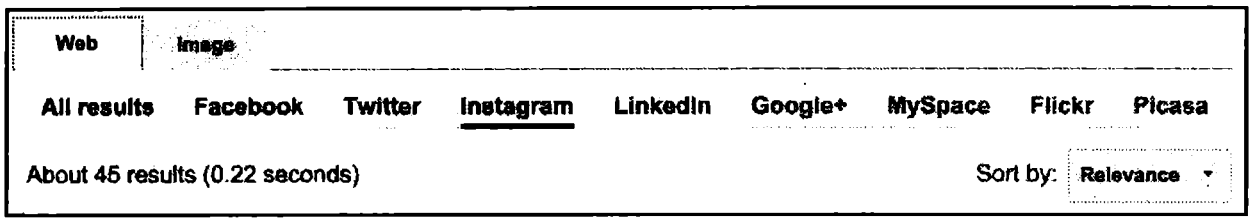


Figure 14.03: A Google CSE that filters images by network.

Exif Data

Every digital photograph captured with a digital camera possesses metadata known as Exif data. This is a layer of code that provides information about the photo and camera. All digital cameras write this data to each image, but the amount and type of data can vary. This data, which is embedded into each photo "behind the scenes", is not visible by viewing the captured image. You need an Exif reader, which can be found on websites and within applications. Keep in mind that some websites remove or "scrub" this data before being stored on their servers. Facebook, for example, removes the data while Twitter and Flickr often do not. Locating a digital photo online will not always present this data. If you locate an image that appears full size and uncompressed, you will likely still have the data intact. If the image has been compressed to a smaller file size, this data is often lost. Any images removed directly from a digital camera card will always have the data. This is one of the reasons you will always want to identify the largest version of an image when searching online. A software application to identify this data will be discussed later. The easiest way to see the information is through an online viewer.

Jeffrey's Exif Viewer (exif.regex.info/exif.cgi)

I consider Jeffrey's Exif Viewer the online standard for displaying Exif data. The site will allow analysis of any image found online or stored on a drive connected to your computer. The home page provides two search options. The first allows you to copy and paste an address of an image online for analysis. Clicking "browse" on the second option will open a file explorer window that

will allow you to select a file on your computer for analysis. The file types supported are also identified on this page. The first section of the results will usually provide the make and model of the camera used to capture the image. Many cameras will also identify the lens used, exposure settings, flash usage, date and time of capture, and file size. In one example, I could see that the camera used was a Canon EOS Digital Rebel with an 18 - 55mm lens at full 55mm setting. Auto exposure was selected, the flash was turned off, and the photo was taken at 2:30 pm on May 7, 2011. Not all of this will be vital for the researcher, but every bit of intelligence counts.

Scrolling down the analysis page will then identify many camera settings that probably provide little information to the researcher. These include aperture information, exposure time, sharpness, saturation, and other image details. Mixed in with this data is the serial number field. This is most common in newer SLR cameras and will not be present in less expensive cameras. These cameras usually identify the make, model, and serial number of the camera inside every photo that they capture. A serial number of a camera associated with an image can be valuable data. This can help an analyst associate other photos found with a target's camera. If an "anonymous" image was found online that included a serial number in the Exif data, and another image was found of a target of the investigation, these two photos can be analyzed. If the serial number as well as make and model of camera match, there is a good likelihood that the same camera took both images. It is important to know that this data can be manipulated, though. Using software such as Exif Tool (see Chapter Nineteen), a user can modify this data. While this is not a popular tactic to use, it is still possible. The difficult part of this is finding photos only knowing the serial number. Two services will help with that.

Stolen Camera Finder (www.stolencamerafinder.co.uk)

This site was designed to help camera theft victims with locating their camera if it is being used by the thief online. For that use, you would find a photo taken with the stolen camera, and drop it into the site for analysis. This analysis identifies a serial number if possible. If one is located, the service then presents links to photo-sharing websites, such as Flickr, that contain photos with the same serial number. This is very creative, and we can use the same service for our own research.

On the home page, click the "no photo?" link in the lower center portion. This will now allow for a manual input of a serial number. Clicking the "page" link will present you with options for payment. In 2013, this site became a premium service and full use is no longer free. However, we can still extract relevant data. Clicking the image thumbnail will open a new window with a compressed version of the image that was located. Conducting a reverse image search, which will be explained in just a moment, will often identify the original copies of this photo that appear online. This will often identify the photographer's screen name which can lead to more information. It is important to verify the make and model number as well, since different camera manufacturers may use overlapping serial numbers.

Camera Trace (cameratrace.com/trace)

An additional site, which is still free, that provides this service is called Camera Trace. Type in the serial number of a camera and the site will attempt to locate any online photographs taken with the camera. This service claims to have indexed all of Flickr and 500px with plans to soon add Smugmug, Picasa, and Photobucket. A sample search using a serial number of "123" revealed results for Panoramio, even though the site is not listed as being included in the results. The website urges users to sign up for a premium service that will make contact if any more images appear in the database. The fee for this is $10 per camera. The website at cameratrace.com/law-enforcement states that the service is free for law enforcement.

GPS

Many new SLR cameras, and almost all cellular telephone cameras, now include GPS. If the GPS is on, and the user did not disable geo-tagging of the photos in the camera settings, you will get location data within the Exif data of the photo. Figure 14.04 (left) displays the analysis of an image taken with a camera with GPS. The data is similar to the previous analysis, but includes a new "Location" field. This field will translate the captured GPS coordinates from the photo and identify the location of the photo. Further down this results page, the site will display an image from Google Maps identifying the exact point of the GPS associated with the photo. Figure 14.04 (right) displays this satellite view including a direction identifier. Since most cellular telephones possess an accelerometer, the device documents the direction the camera was facing. Most Android and iPhone devices have this capability. Your results will vary depending on the user's configuration of their GPS on the device.

Camera:	Apple iPhone 5
Lens:	4.1 mm
Exposure:	Auto exposure, Program AE, 1/452 sec, f/2.4, ISO 50
Flash:	Off, Did not fire
Date:	**March 24, 2013** 5:13:57PM (timezone not specified) (1 year, 10 months, 13 days, 18 hours, 20 minutes, 25 seconds ago, assuming image timezone of 6 hours behind GMT)
Location:	Latitude/longitude: **38°** 30' 43.8" North, **90°** 30' 0" West (38.512167, -90.500000) Location guessed from coordinates: *1526 Atlantic Crossing Drive, Fenton, MO 63026, USA*



Figure 14.04: A Jeffrey's Exif Viewer result identifying location with map view.

Cropped Images

Another piece of information that we can look for from the Exif data is the presence of a thumbnail image within the photograph. Digital cameras generate a small version of the photo captured and store it within the Exif data. This icon size image adds very little size to the overall file. When a user crops the image, this original smaller version may or may not get overwritten.

Programs such as Photoshop or Microsoft Photo Editor will overwrite the data and keep both images identical. Other programs, as well as some online cropping tools, do not overwrite this data. The result is the presence of the original and uncropped image within the Exif data of the cropped photo. An example of this is seen in Figure 14.05. A cropped photo found online is examined through Jeffrey's Exif viewer. The cropped full size large photo is seen on the left. The embedded smaller original photo was not overwritten when cropped. We can now see what the image looked like before it was cropped. This technique has been used by police to identify child pornography manufactures. These pedophiles will crop themselves out of illegal images to avoid identification. When photos of the children are found by police, an original uncropped image may be enough to identify and prosecute a molester. This is not limited to law enforcement. Some tech savvy fans of television personality Catherine Schwartz examined a cropped photo on her blog in 2003. Inside the Exif data was the un-cropped version which exposed her breasts and quickly made the rounds through the internet.

Real World Application: In a civil litigation, a subject claimed an injury that prohibited him from work, walking, and a normal life. The suit claimed damages from pain and suffering and sought a monetary judgment for future lack of ability to work. A brief scan of the subject's online photo album revealed fishing trips, softball games, and family adventure vacations. With Exif information data intact, exact dates, times, locations, and cameras were identified and preserved. The subject withdrew his lawsuit.

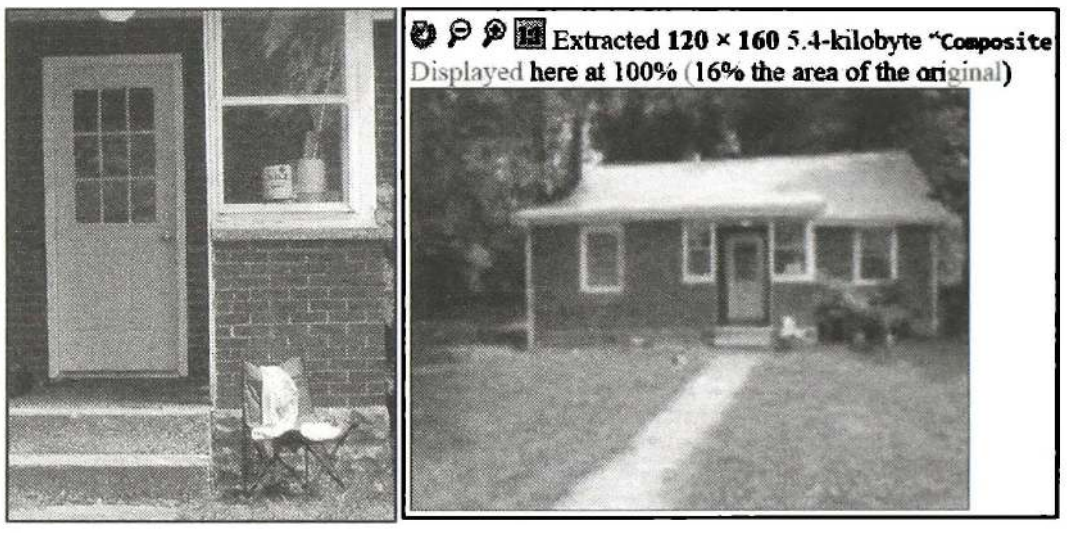


Figure 14.05: A Jeffrey's Exif Viewer summary result displaying an original uncropped photo.

Exif Search (exif-search.com)

The idea of searching within the metadata stored within images has intrigued online researchers for years. Many people are surprised that Google has not tackled this obstacle. Exif Search is a service that allows limited search within this metadata, but it is far from perfect. After creating a free account, the main page will present a single search field. Here, you can search by date range, serial number, manufacturer, or any keywords. The free trial limits you to ten searches. While I have had little success with this service, its usage should be considered when needed.

Online Barcode Reader (online-barcode-reader.inliteresearch.com)

Barcodes have been around for decades. They are the vertical lined images printed on various products that allow registers to identify the product and price. Today's barcodes are much more advanced and can contain a large amount of text data within a small image. Some newer barcodes exist in order to allow individuals to scan them with a cell phone. The images can provide a link to a website, instructions for downloading a program, or a secret text message. I generally advise against scanning any unknown barcodes with a mobile device since malicious links could be opened unknowingly. However, an online barcode reader can be used to identify what information is hiding behind these interesting images. Figure 14.06 displays the barcode search options from Online Barcode Reader. These include 1D, PDF417, Postal, DataMatrix, QR, and ID barcodes. After selecting the type of barcode image, you can select any PDF, TIFF, JPEG, BMP, GIF, or PNG file on your computer up to 4Mb in size. This could be a photo that possesses a barcode in the content or a digital code downloaded from a website. Screen captures of codes also work well. While sitting on a plane with Wi-Fi, I captured a photo of an abandoned boarding pass in the magazine holder in front of me. The barcode reader identified text information stored inside the code that was not present in text on the document.

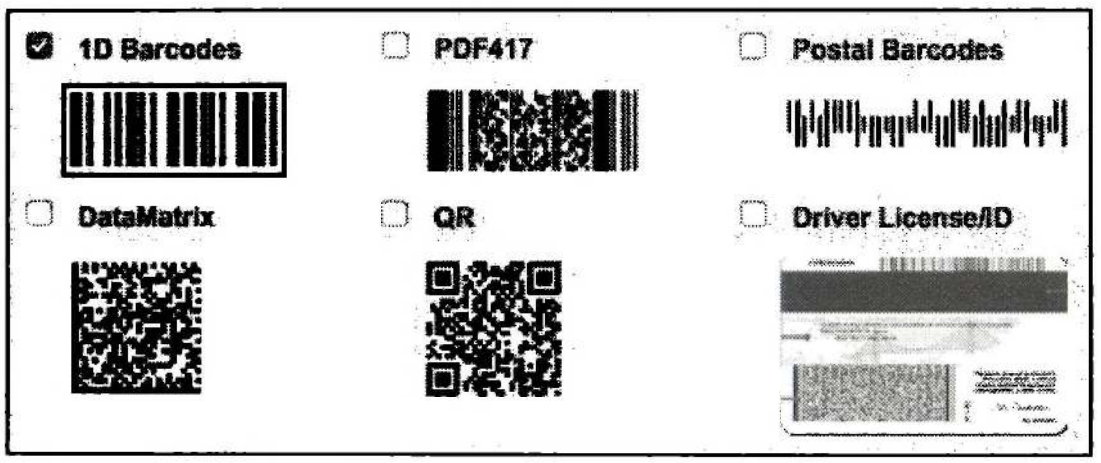


Figure 14.06: Barcode input samples from Online Barcode Reader.

Image Manipulation

It is common to find images on the internet that have been manipulated using software such as Photoshop. Often it is difficult, if not impossible, to tell if these photos have been manipulated by visually analyzing them. A handful of websites use a technique to determine not only if the photo has been manipulated, but which portions of the photo have changed. One site offers the following explanation of how the technology works.

"Error level analysis (ELA) works by intentionally resaving the image at a known error rate, such as 95%, and then computing the difference between the images. If there is virtually no change, then the cell has reached its local minima for error at that quality level. However, if there is a large amount of change, then the pixels are not at their local minima and are effectively original."

Foto Forensics (fotoforensics.com)

This site allows you to upload a digital image. After successful upload, it will display the image in normal view. Below this image will be a darkened duplicate image. Any highlighted areas of the image indicate a possible manipulation. While this site should never be used to definitively state that an image is untouched or manipulated, investigators may want to conduct an analysis for intelligence purposes only. Figure 14.07 displays original and manipulated images while Figure 14.08 displays the analysis of the images from Foto Forensics. This site will provide an analysis of an image from the internet or a file uploaded from a computer. It is important to note that any images uploaded become part of the website's collection and a direct URL is issued. While it would be difficult for someone to locate the URL of the images, it could still pose a security risk for sensitive photographs.

Izitru (izitru.com)

The previous service will identify the areas of a photo that have been manipulated. If the edited content is not obvious, you may want a service that will provide additional analysis. Izitru may provide the details that you need in an investigation. I conducted a test with an original photo from a cellular telephone camera. Figure 14.09 (above) displays the result from the unmodified image. Figure 14.09 (below) displays the result after I modified a very small portion of the image. Identifying images that have been re-saved instead of copied could be valuable in civil litigation.

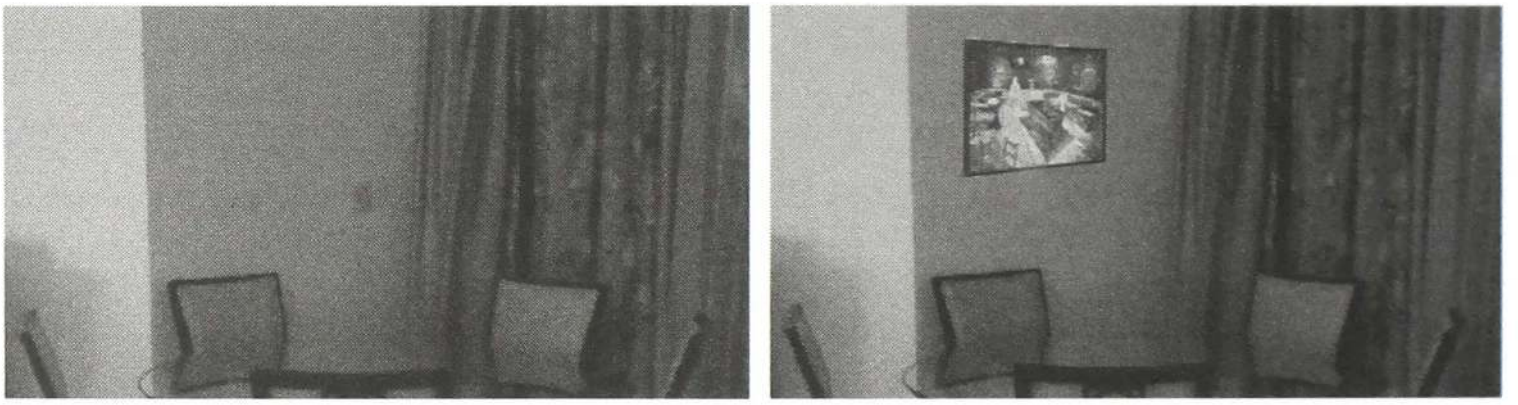


Figure 14.07: An original photograph (left) compared to a manipulated photograph (right).

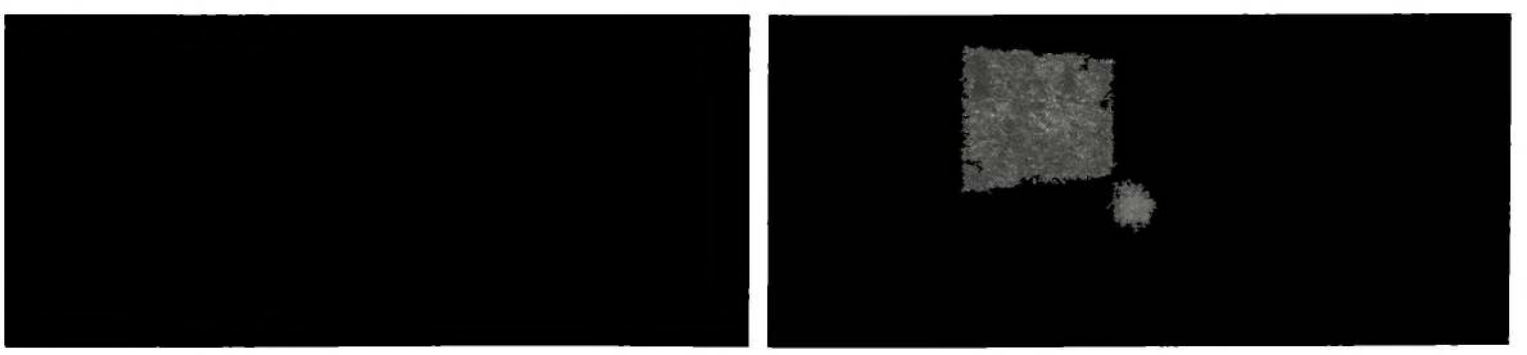


Figure 14.08: The original photograph (left) and manipulated image (right) on Foto Forensics.

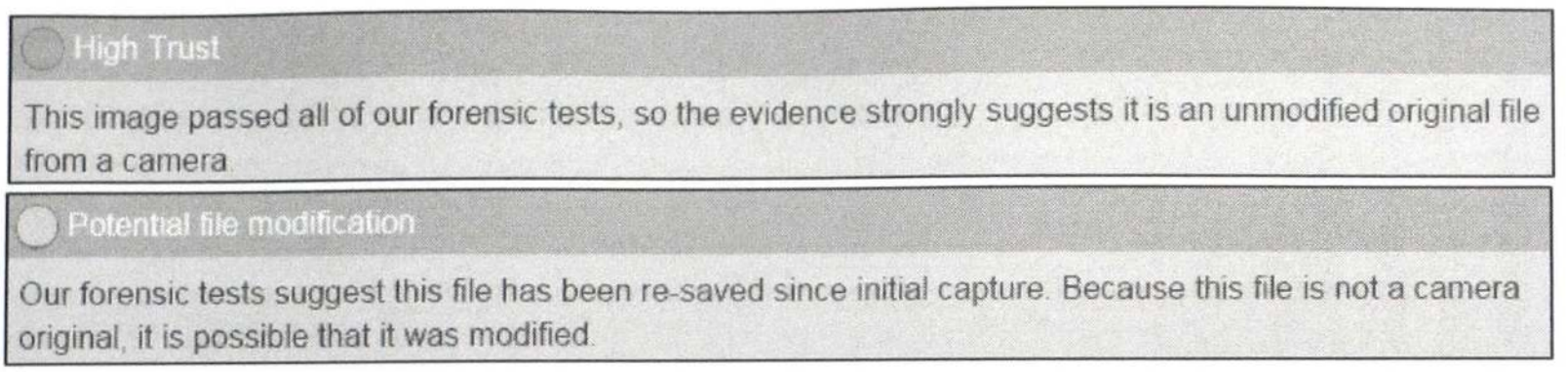


Figure 14.09: Results from Izitru of an unmodified (above) and manipulated image (below).

Forensically (29a.ch/photo-forensics)

Forensically is a robust image analyzer that offers a huge collection of photo forensic tools that can be applied to any uploaded image. This type of analysis can be vital when image manipulation is suspected. Previous tools have offered one or two of the services that Forensically offers, but this new option is an all-in-one solution for image analysis. Loading the page will present a demo image, which is used for this explanation. Clicking the "Open File" link on the upper left will allow upload of an image into your browser for analysis. Images are NOT uploaded to the server of this tool, they are only brought into your browser locally. Figure 14.10 (left) is the standard view of a digital photo. The various options within Forensically are each explained and example images are included. Due to the black & white environment of this book, I have replicated all of this instruction in color on my blog at the following address.

https://inteltechniques.com/wp/2016/12/21/internet-search-resource-foresically/

The Magnifier allows you to see small hidden details in an image. It does this by magnifying the size of the pixels and the contrast within the window. There are three different enhancements available at the moment: Histogram Equalization, Auto Contrast, and Auto Contrast by Channel. Auto Contrast mostly keeps the colors intact; the others can cause color shifts. Histogram Equalization is the most robust option. You can also set this to none.

The Clone Detector highlights copied regions within an image. These can be a good indicator that a picture has been manipulated. Minimal Similarity determines how similar the cloned pixels need to be to the original. Minimal Detail controls how much detail an area needs; therefore, blocks with less detail than this are not considered when searching for clones. Minimal Cluster Size determines how many clones of a similar region need to be found in order for them to show up as results. Blocksize determines how big the blocks used for the clone detection are. You generally don't want to touch this. Maximal Image Size is the maximal width or height of the image used to perform the clone search. Bigger images take longer to analyze. Show Quantized Image shows the image after it has been compressed. This can be useful to tweak Minimal Similarity and Minimal Detail. Blocks that have been rejected because they do not have enough detail show up as black. Figure 14.10 (right) demonstrates this output.

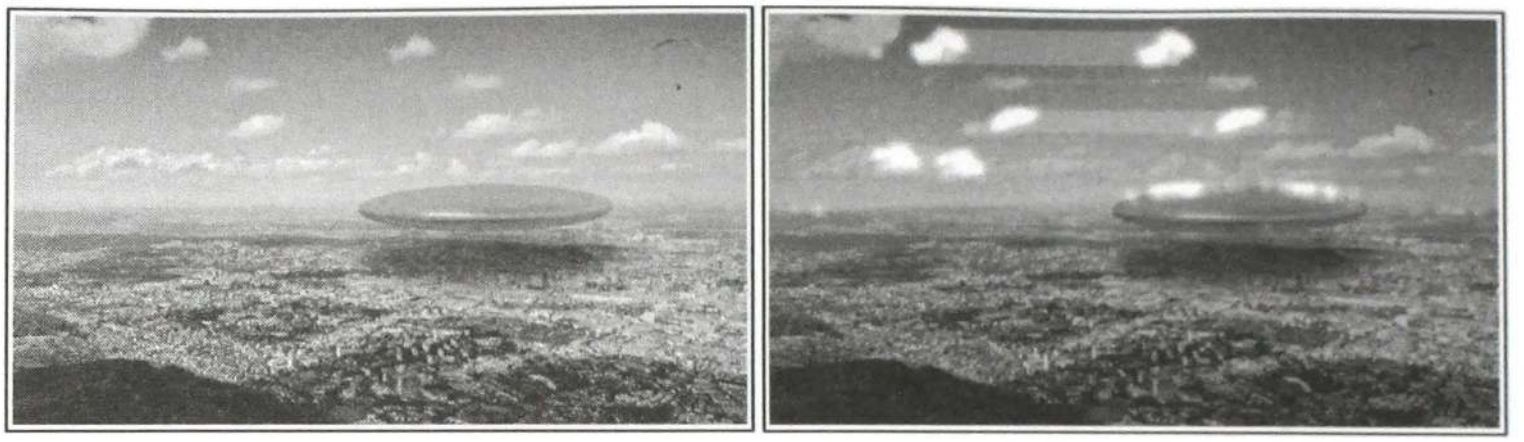


Figure 14.10: A normal image view (left) and Clone Detector (right) in Forensically.

Error Level Analysis compares the original image to a recompressed version. This can make manipulated regions stand out in various ways. For example, they can be darker or brighter than similar regions which have not been manipulated. JPEG Quality should match the original quality of the image that has been photoshopped. Error Scale makes the differences between the original and the recompressed image bigger. Magnifier Enhancement offers different enhancements: Histogram Equalization, Auto Contrast, and Auto Contrast by Channel. Auto Contrast mostly keeps the colors intact; the others can cause color shifts. Histogram Equalization is the most robust option. You can also set this to none. Opacity displays the opacity of the differences layer. If you lower it you will see more of the original image. Figure 14.11 (left) displays manipulation.

Noise Analysis is basically a reverse de-noising algorithm. Rather than removing the noise it removes the rest of the image. It is using a super simple separable median filter to isolate the noise. It can be useful for identifying manipulations to the image like airbrushing, deformations, warping, and perspective corrected cloning. It works best on high quality images. Smaller images tend to contain too little information for this to work. Noise Amplitude makes the noise brighter. Equalize Histogram applies histogram equalization to the noise. This can reveal things but it can also hide them. You should try both histogram equalization and scale to analyze the noise. Magnifier Enhancement offers three different enhancements: Histogram Equalization, Auto Contrast, and Auto Contrast by Channel. Auto Contrast mostly keeps the colors intact; the others can cause color shifts. Histogram Equalization is the most robust option. You can also set this to none. Opacity is the opacity of the noise layer. If you lower it you will see more of the original image. The result can be seen in Figure 14.11 (right).

Level Sweep allows you to quickly sweep through the histogram of an image. It magnifies the contrast of certain brightness levels. To use this tool simply move your mouse over the image and scroll with your mouse wheel. Look for interesting discontinuities in the image. Sweep is the position in the histogram to be inspected. You can quickly change this parameter by using the mouse wheel while hovering over the image, this allows you to sweep through the histogram. Width is the amount of values (or width of the slice of the histogram) to be inspected. The default should be fine. Opacity is the opacity of the sweep layer. If you lower it you will see more of the original image.

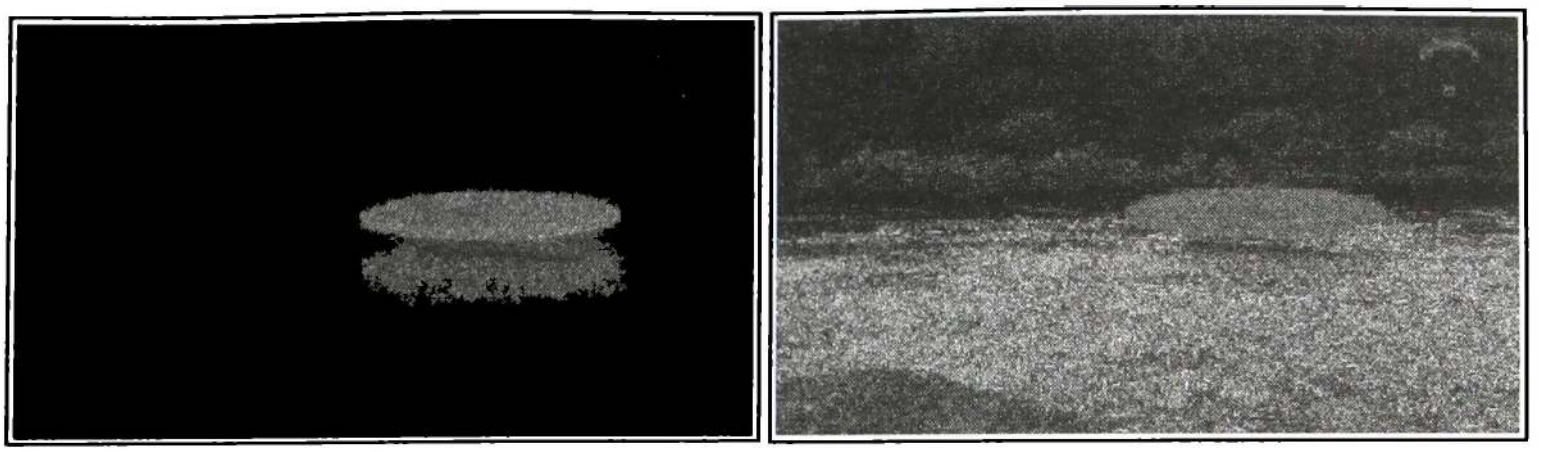


Figure 14.11: Error Level Analysis (left) and Noise Analysis (right) in Forensically.

Luminance Gradient analyzes the changes in brightness along the x and y axis of the image. Its obvious use is to look at how different parts of the image are illuminated in order to find anomalies. Parts of the image which are at a similar angle (to the light source) and under similar illumination should have a similar color. Another use is to check edges. Similar edges should have similar gradients. If the gradients at one edge are significantly sharper than the rest it's a sign that the image could have been copied and pasted. It does also reveal noise and compression artifacts quite well. Figure 14.12 (left) displays this view.

PCA performs principal component analysis on the image. This provides a different angle to view the image data which makes discovering certain manipulations and details easier. This tool is currently single threaded and quite slow when running on big images. Choose a Mode: Projection of the value in the image onto the principal component; Difference between the input and the closest point on the selected principal component; Distance between the input and the closest point on the selected principal component; or the closest point on the selected principal Component. There are three different enhancements available: Histogram Equalization, Auto Contrast, and Auto Contrast by Channel. Auto Contrast mostly keeps the colors intact; the others can cause color shifts. Histogram Equalization is the most robust option. You can also set this to none. Opacity is the opacity of the sweep layer. If you lower it you will see more of the original image. Figure 14.12 (right) displays this view.

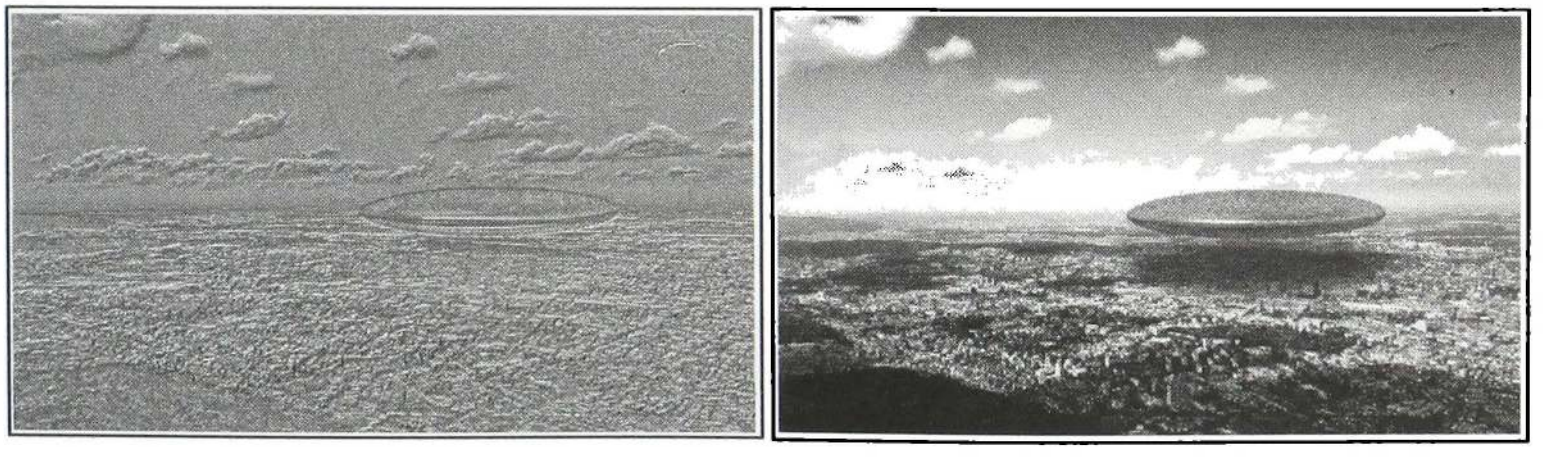


Figure 14.12: The Luminance analysis (left) and PCA analysis (right) within Forensically.

MetaData displays any Exif metadata in the image. Geo Tags shows the GPS location where the image was taken, if it is stored in the image. Figure 14.13 displays the result.

Thumbnail Analysis shows any hidden preview image inside of the original image. The preview can reveal details of the original image or the camera used. Figure 14.14 displays the online image (left) while the original thumbnail displays a different view (right).

The next time you identify a digital image as part of your online investigation, these tools will peek behind the scenes and may display evidence of tampering.

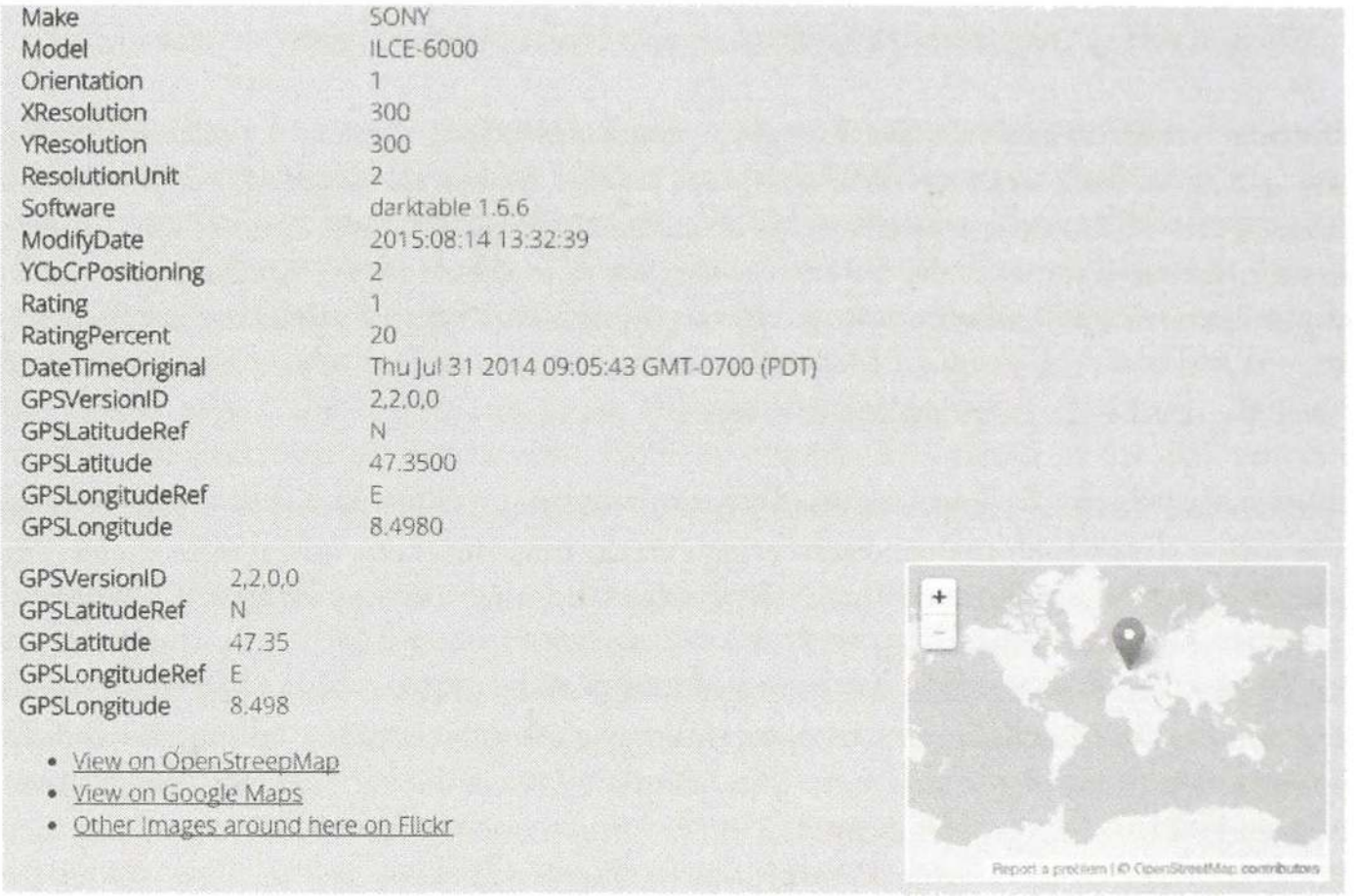


Figure 14.13: Metadata from Forensically.

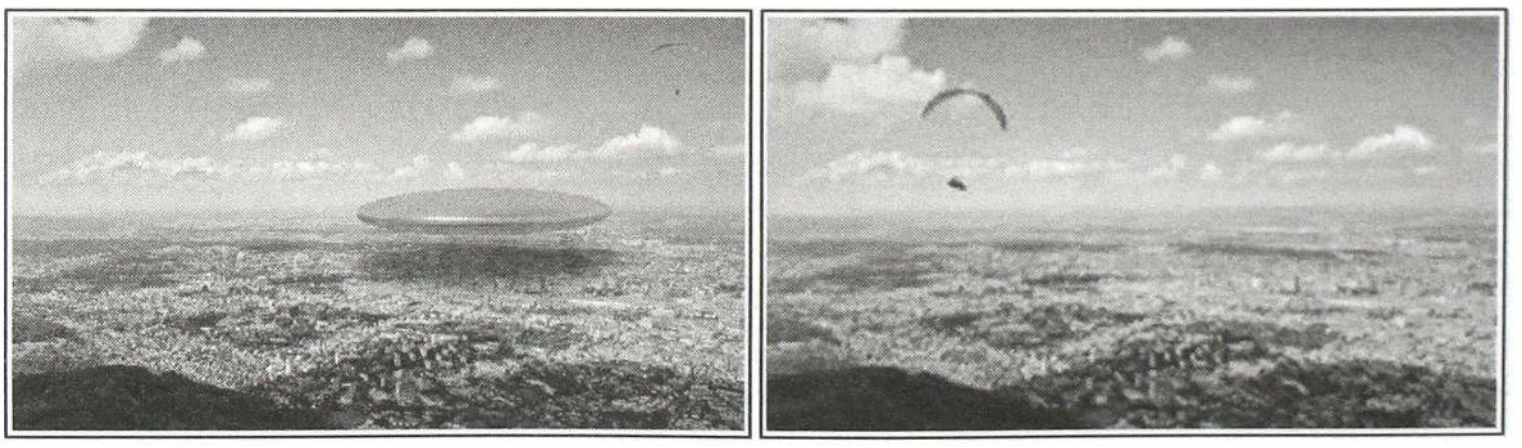


Figure 14.14: An online image (left) and original thumbnail image (right) on Forensically.

CHAPTER FIFTEEN
VIDEOS

Online videos are almost as common as online photographs. The cameras in data phones can act as video cameras. In some situations, uploading a video to the internet is easier than a photograph. Video sharing sites such as YouTube have made video publication effortless. For investigations, a video can contain a huge amount of intelligence. When any abnormal event happens, people flock to their phones and start recording. These videos may capture criminal acts, embarrassing behavior, or evidence to be used in a civil lawsuit. Obtaining these videos is even easier than creating them.

YouTube (YouTube.com)

The most popular of several video-sharing sites is YouTube. The official YouTube site declares that 48 hours of video are uploaded every minute, resulting in nearly 8 years of content uploaded every day. It further states that over 4 billion videos are viewed each day. These impressive statistics confirm the need to include videos as part of a complete OSINT analysis. YouTube is easy to search from the main search field on every page. This field can accept any search term and will identify video content or user name. Users that upload videos to YouTube have their own "channel". Their videos are uploaded to this channel, and locating a user's channel will identify the videos uploaded by that user.

Many people use YouTube as a social network, leaving comments about videos and participating in discussions about various topics. If you locate a video of interest, it is important to also retrieve this text information. Each comment below a video will include the user name that created the comment, which will link to that user's profile.

A search for "school bus fight" returned over 342,000 video links on YouTube. Adding a search term such as the city or school name may help, but it may also prohibit several wanted videos from appearing. The "filters" option can be expanded to help limit the search scope. This button is above the first video result. This provides additional filter options including the ability to sort by the upload date (date range), type (video vs. channel), duration (short or long), and features (video quality). In the "school bus fight" example, the "uploaded this week" option was chosen. This resulted in only 437 videos which could easily be examined for any intelligence. The lower left portion of any video page includes a link to the profile of the user who submitted this video. This profile page includes all of the videos uploaded by that user and additional profile information. Several YouTube "hacks" have surfaced over the years. Many of these stopped working as YouTube made changes to the environment. Of those still functioning, I find the following techniques helpful to my investigations.

Bypass Age and Login Restriction

Several YouTube videos have been tagged as violent, sexual, or otherwise inappropriate for young viewers. Others demand that you log into a Google account in order to view the content for unclear reasons. Either way, this is an unnecessary roadblock to your investigation. As an OSINT investigator, I prefer to not be logged into any personal or covert account while I am researching. Any time you are searching through a Google product while logged into an account, Google is documenting your every move. This can be unsettling. One easy technique should remove this restriction. Navigate to the following website and notice the inability to view the video. If you are not logged into a Google account with a verified age, you should see a warning about mature content. This video cannot be played.

https://www.youtube.com/watch?v=SZqNKAd_gTw

In this example, the YouTube Video ID is "SZqNKAd_gTw". In order to view this video through YouTube without a third-party service, and without supplying the credentials for your personal Google account, you can generate the following URL. Replace "SZqNKAd_gTw" with the ID of your target video. The result will be the restricted video in full screen view. Some users report that this technique will also bypass videos that have a viewing country restriction.

https://www.youtube.com/v/SZqNKAd_gTw

If this technique should ever stop working, you can also use a non-Google service to achieve the same result. Navigate to the following website to access this same video (SZqNKAd_gTw).

http://www.nsfwyoutube.com/watch?v= SZqNKAd_gTw

Notice that all of the addresses are very similar. This final link will take you to NSFWYouTube, a third-party website, which will also remove the proof of age requirement. Please be warned that the content in this example contains very disturbing video, hence the blockage by YouTube.

Bypass Commercials with Full Screen

It seems lately that every long YouTube video I play possesses a 30 second commercial at the beginning. This is very frustrating when analyzing a large number of videos. A quick URL trick will bypass this annoyance. Navigate to the following address and notice the long commercial at the beginning.

http://www.youtube.com/watch?v=IEIWdEDFlQY

Alter this address slightly in order to force the video to play in full screen in your browser. This will also bypass any commercials. The URL should appear like the following.

https://youtube.googleapis.com/v/IEIWdEDFlQY

Display Frames of Videos

When a user uploads a video, YouTube captures four image frames that are publicly visible if a proper query is submitted. Navigate to the following address to load an example video.

https://www.youtube.com/watch?v=gsnmUdGnJhc

Using that same video ID, navigate to the following address to view the main still frame. This is the image visible when a video is loaded within YouTube before playing.

http://i.ytimg.com/vi/OmZyrynlk2w/0.jpg

The address that displayed the main image is not your only option. At least four images can be extracted from this specific video with the following addresses. Each of these could be searched with the reverse image techniques explained in the previous chapter.

http://i.ytimg.com/vi/OmZyrynlk2w/0.jpg
http://i.ytimg.com/vi/OmZyrynlk2w/1.jpg
http://i.ytimg.com/vi/OmZyrynlk2w/2.jpg
http://i.ytimg.com/vi/OmZyrynlk2w/3.jpg

Identify and Bypass Country Restriction

Many videos on YouTube are allowed to be viewed in some countries and blocked in others. If you encounter a video that will not play for you because of a country restriction, you have options. Before proceeding, consider identifying from which geographical areas a video is restricted. For this, we will rely on the polsy.org.uk country restriction checker. After you have identified a video with possible country restrictions, paste the video ID into the following URL. Note that the following video ID (MGhMdT_C-vQ) is blocked in the U.S.

http://polsy.org.uk/stuff/ytrestrict.cgi?ytid=MGhMdT_C-vQ

The result is a page with a world map. Countries in grey are allowed to view the target video while countries in red are not. While I cannot natively play this video due to my location, I can easily view the four still frames with the technique described in the previous section. The following exact URLs display content otherwise not viewable.

http://i.ytimg.com/vi/MGhMdT_C-vQ/0.jpg
http://i.ytimg.com/vi/MGhMdT_C-vQ/1.jpg
http://i.ytimg.com/vi/MGhMdT_C-vQ/2.jpg
http://i.ytimg.com/vi/MGhMdT_C-vQ/3.jpg

If a video is blocked from playing in your location, you can use a third-party proxy that should allow viewing. In both of these examples, MGhMdT_C-vQ is the target video that is blocked. The following URL should play the video.

https://www.youpak.com/watch?v=MGhMdT_C-vQ

An alternative to this is HookTube, which would be the following URL for this video.

https://hooktube.com/watch?v= MGhMdT_C-vQ

The URL below will not stream the video, but should allow you to download the content.

https://www.ssyoutube.com/watch?v=MGhMdT_C-vQ

Metadata and Reverse Image Search (www.amnestyusa.org/citizenevidence)

Most of the details of a YouTube video can be seen on the native page where the video is stored. Occasionally, some of this data may not be visible due to privacy settings or profile personalization. In order to confirm that you are retrieving all possible information, you should research the data visible from YouTube's servers. The easiest way to do this is through the YouTube Data Viewer from Amnesty International. Select any YouTube video and copy the entire address of the page. Paste the video address into this service. The result will include the video title, description, upload date, upload time, four still images, and an option to conduct a reverse image search. Clicking the "reverse image search" option next to a still frame opens a reverse image search on Google for the selected image. While this works well on YouTube videos, complete reverse video searching across multiple networks will be explained later in this chapter.

Immediate Download Options

My preferred method for extracting YouTube videos was explained in Chapter Two while discussing Youtube-DL. However, if you have no software or browser plugins available to you, there is another easy option. While you are watching any YouTube video, you can add the letters "PWN" to the beginning of the address in order to download the video to your computer. To test this, navigate to http://www.youtube.com/watch?v= OmZyrynlk2w.

Now, add "PWN to the beginning, as indicated in the following address.

http://www.pwnyoutube.com/watch?v= OmZyrynlk2w

You will be presented a new page with many options including the ability to download the video; download only the audio; convert the video to a different format; and bypass the age restriction as discussed earlier. Additional options include yout.com, keepvid.com, and y2mate.com.

YouTube Comment Scraper (http://ytcomments.klostermann.ca/scrape)

While the video media of your target is essential to your investigation, the commentary associated with the posted content is equally important. Screen captures can usually identify the bulk of the comments attached to a video, but there are hurdles. Some viral videos will have thousands of comments, which will not appropriately fit into a screen capture. Other videos may have comments to comments, and those may need to be expanded manually in order to observe and collect the evidence. The solution to this is the YouTube Comment Scraper. This website allows easy use of open-source code available to extract YouTube comments from your own server. Simply provide the URL of your target video and allow the tool to process all data. The result will be a CSV spreadsheet with every comment text, including comment ID, timestamp, user name, replies, and likes. This can all be viewed within the results screen or downloaded (my preference). Figure 15.01 displays a partial collection of captured data.

user	date	timestamp	commentText
Alois Hebenstreit	2 days ago	1514316998089	music is crazzy
James Lee	1 week ago	1513884998091	Finally a band that speaks the truth <3 Billy Talent
Yukio Nana	1 week ago	1513884998094	Tobito REEEEEEEEEEEEEEEEEEEEEEEE
GERHARDTJEAN	1 week ago	1513884998096	Chapado e escutando isso!!! Mas é bom demais!!!!

Figure 15.01: Data collected by YouTube Comment Scraper.

I believe that this type of documentation should be a part of every investigation associated with a video. If you ever find yourself downloading a target video, immediately jump to this free resource and collect the comments. If the video is actively discussed, you may need to return often and download the comments. If you use a program such as Excel, it will be easy to remove duplicates from all of your captures, leaving only the unique comments relative to your investigations.

YouTube Geo Search (https://youtube.github.io/geo-search-tool/search.html)

Many YouTube users intentionally tag their uploads with location data. This online tool allows entry of a location name, address, or set of GPS coordinates. The results identify YouTube videos tagged within the immediate area. Zooming of the interactive map is supported, but the reliability of the exact locations is mediocre. Since this data is user-generated, and not collected via the location of the device, you can only rely on the location set by the target. On many of my investigations, I was only given the city of the video, and nothing further. While this tool can provide videos of interest based on a location, it has rarely assisted with identification of a YouTube user name or video.

IntelTechniques YouTube Search Tool (inteltechniques.com/osint/youtube.html)

If you feel overwhelmed at this point, I understand. I found myself getting confused about which address was the most appropriate for each technique. I created a web page that will walk you through the process. Navigate to the above address to access an all-in-one option. This page will allow you to enter only the user ID of each video of interest. Embedded JavaScript will formulate and execute each address for you through your own browser, and not on my server. This should simplify the processes that were explained previously. These options will take a video ID and provide a full screen view with commercial bypass, age restriction bypass, and country restriction bypass. It will identify countries where a video is restricted and download options for the target video. It will display the four still frames of a target and conduct the entire metadata view of any YouTube video within this page. Figure 15.02 displays the current state of this tool.

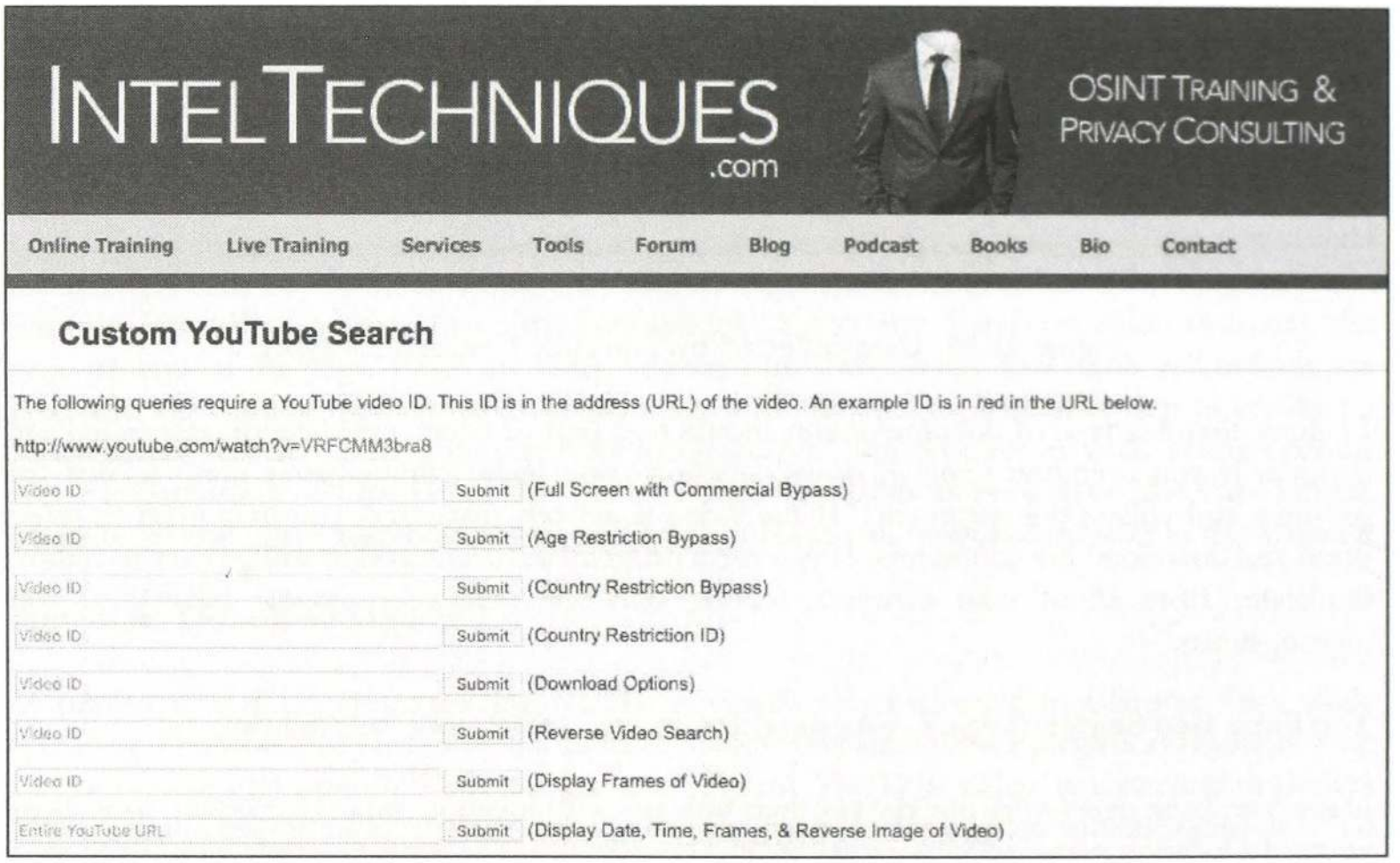


Figure 15.02: The IntelTechniques Custom YouTube Search Tool.

Reverse Video Searching

There was a brief mention earlier of conducting a reverse image search on the four still captures of a YouTube video. This would use the same techniques as mentioned in Chapter Fourteen for images. While there is no official reverse video search option, applying the techniques to still captures of videos can provide amazing results. This method is not limited to YouTube. We can conduct reverse image searches on videos from many sites. As the popularity of online videos is

catching up to images, we must always consider reverse video searches. They will identify additional websites hosting the target videos of interest. Before explaining the techniques, consider the reasons that you may want to conduct this type of activity.

School resource officers and personnel are constantly notified of inappropriate video material being posted online. Identifying these videos, consisting of fights, malicious plans, and bullying, may be enough to take care of the situation. However, identifying the numerous copies on other websites will help understand the magnitude of the situation.

Global security divisions monitoring threats from protest groups will likely encounter videos related to these activities. However, identifying the numerous copies on various websites will often disclose commentary, online discussions, and additional threats not seen in the primary source video.

Human trafficking investigators can now reverse search videos posted to escort providers such as Backpage. The presence of identical videos posted to numerous geographical areas will highlight the travel and intentions of the pimps that are broadcasting the details.

There are limitless reasons why reverse video searching should be a part of your everyday research. The following methods will get you started with the most popular video websites. These techniques can be replicated as you find new services of interest. At the end, I share my custom reverse video search tool that will automate the entire process.

YouTube: As explained earlier, YouTube offers four still frames for every video uploaded. Obtain the URLs outlined during that instruction, and provide each to Google, Bing, Yandex, TinEye, and Baidu as a reverse image search as explained in Chapter Fourteen.

Vimeo: Vimeo does not natively offer URLs with a video ID that display screen captures of multiple frames. However, they do provide a single high definition still capture for every video. This is stored in the Application Programming Interface (API) side of Vimeo, but it is easy to obtain. As an example, consider your target is at https://vimeo.com/99199734. The unique ID of 99199734 is assigned to that video. You can use that number to access the API view of the video at https://vimeo.com/api/oembed.json?url=https://vimeo.com/99199734. This address will display a text only page with the following content.

type: "video",
version: "1.0",
provider_name: "Vimeo",
provider_url: "https://vimeo.com/",
title: "Billy Talent 'Try Honesty'",
author_name: "S M T",
author_url: "https://vimeo.com/user10256640",
is_plus: "0",

html: "<iframe src="https://player.vimeo.com/video/99199734" width="480" height="272" frameborder="0" title="Billy Talent 'Try Honesty'" webkitallowfullscreen mozallowfullscreen allowfullscreen></iframe>",
width: 480,
height: 272,
duration: 247,
description: "Music Video Directed by Sean Michael Turrell. *Winner MuchMusic Awards Best Rock Video.",
thumbnail_url: "https://i.vimeocdn.com/video/513053154_295x166.jpg",
thumbnail_width: 295,
thumbnail_height: 167,
upload_date: "2014-06-25 21:29:06",
video_id: 99199734,
uri: "/videos/99199734"

The portion relevant to this topic is the thumbnail URL. Following that description is an exact address of the large image used at the beginning of each Vimeo video. Also note that this view identifies the exact date and time of upload. The video page view only identified the date as "1 Year Ago". A reverse image search of the thumbnail URL will produce additional websites that host the same (or similar) video.

Facebook: Identifying Facebook screen captures does not require access to their API. Instead, we will view the source code of an individual video page. For this example, assume that you are viewing a video embedded into a Facebook profile. You can right-click the video while playing and choose the option "Show Video URL". This will present a small box with the address of the Facebook video that is playing. In this example, we will use the following.

https://www.facebook.com/billytalent/videos/10153157582551992/

Right-click on this new page and choose to view the source code. Conduct a search of .jpg and copy the entire URL of the first search result. In this example, the page includes the following image address.

https://scontent-lax3-1.xx.fbcdn.net/v/t15.0-10/10604952_10153157584261992_1015315758 2551992_3872_390_b.jpg?oh=1c06c278ac42a6ddd66b41507444f647&oe=5AE94A65

Navigating to that URL presents a large primary screen capture of the target video. Conducting a reverse image search through various services will identify additional websites that possess the same content. These will usually be social network profiles of the target's friends. In this example, a reverse search of this image reveals 43 additional pages of interest. Each contain the target video and comments about the content.

Vine: Identifying Vine screen captures also does not require access to their API. Instead, we will view the source code of an individual video page. Assume that your target Vine video is located at the URL of https://vine.co/v/eUHhwnul2K6. Right-click on the page and view the source code text. Conduct a search of .jpg and copy the entire URL of the first search result. Note that since Vine has stopped accepting new submissions, some videos no longer have a jpeg file associated with the post. In this example, the post includes the following image address.

https://v.cdn.vine.co/r/videos/52B2D72BF811392_4f9b1663eae.0.2.4118360994298.mp4.jpg

Navigating to that URL should present a large primary screen capture of the target video. Conducting a reverse image search through various services will identify additional websites that possess the same content. These will usually be social network profiles of the target's friends.

Instagram: Instagram can host videos in the same way that it possesses images. Similar to Vine, Instagram will require us to view the source code of a target video. In this, you will see a descriptor of meta property="og:image" near the top of the page. The URL that follows this is the address of the primary screen capture for this video. In this example, it is the following address.

https://igcdn-photos-g-a.akamaihd.net/hphotos-ak-xft1/t51.2885-15/e15/12237557_1654671668132782_1953483052_n.jpg

Liveleak: Similar to Instagram, LiveLeak displays a meta property="og:image" line near the top of every video page's source code view. As an example, consider that your target is a bus fight video located at http://www.liveleak.com/view?i=9d4_1447939701. The source code of that page reveals the still image to be located at the following address.

https://cdn.liveleak.com/80281E/ll_a_u/thumbs/2015/Nov/19/d5e7c911321a_sf_5.jpg

Backpage: While the technique for obtaining the still image of a Backpage video is very similar to the others, there is one main difference. Looking at the source code of any page that hosts a video will not reveal any jpg images or an "os:image" descriptor. Backpage stores their images in the .png format, and this image link can be found by searching png in the source code view. The following is an actual result from the page source of a Backpage profile.

video src="http://video.backpage.com/1-5d953c0e3f9ee71d653611.mp4.mp4"
poster="http://videothumb.backpage.com/1-5d953c0e3f9ee71d653611.mp4.mp4-00002.png"

The second URL is a full-size image of the preview frame for this video. A reverse search identifies several additional copies of this video related to sex trafficking.

Others: Repeating this process for every video sharing website can get redundant quickly. Instead, consider the following. Practically every online video possesses a still image that is displayed to represent the video before being played and in search results. This image is likely a

direct link that can be seen in the source code. Providing this image to various reverse image search websites will likely display additional copies of the target video within unknown websites.

IntelTechniques Reverse Video Search Tool (inteltechniques.com/OSINT/reverse.video.html)

Similar to the previous search tools that I host, this page will automate the instructions explained here. This page includes a search field that awaits a YouTube, Vimeo, Facebook, Vine, Instagram, or LiveLeak video ID (Figure 15.03). It will also accept an entire URL from any Backpage profile that contains a video. The entire URL is required due to their page naming scheme. When executed, several new tabs will open with your reverse image searches for that video from Google, Bing, TinEye, Yandex, and Baidu. The final option allows you to enter the entire URL of any image that you locate and the reverse image process is conducted.

I rely on this tool almost every day, and hope that you will also find it beneficial. Much of the code used to create these tools was provided by Justin Seitz at automatingosint.com. His training on applying OSINT techniques through Python scripts has been extremely valuable.

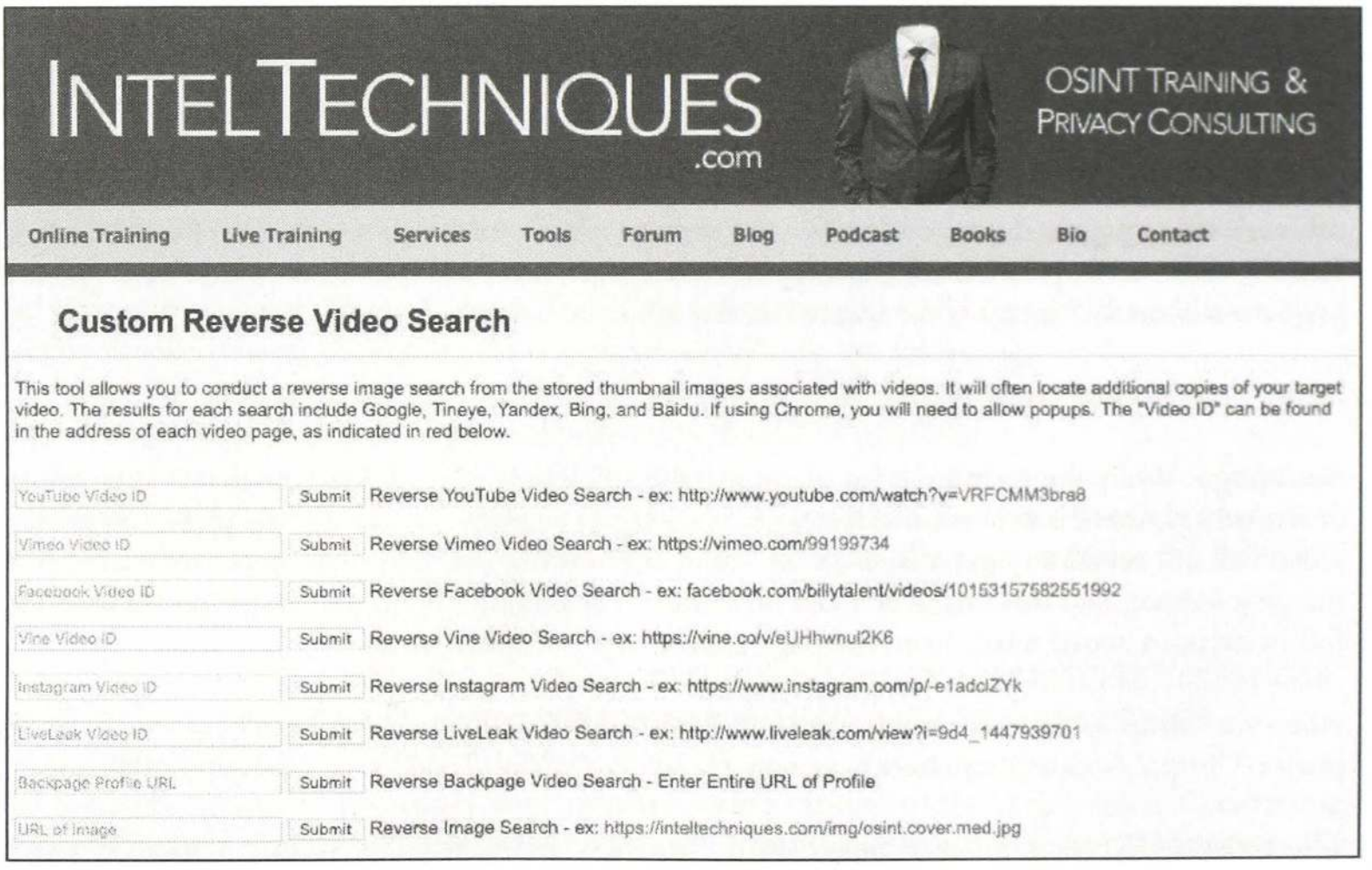


Figure 15.03: The IntelTechniques Custom Reverse Video Search Tool.

These sources are not the only video sharing services on the internet. Wikipedia identifies dozens of these sites, but searching each of them can become tedious. These sites are no longer restricted to pages of video files with standard extensions such as mp4, mpg, and flv. Today, services such as Instagram and Vine allow embedded videos that do not conform to yesterday's video standards. Many new services present the viewer with animated gif files that only appear as true videos. Fortunately, search engines like Google and Bing offer a search across all of the types.

Google Videos (videos.google.com)

A search on YouTube for "school bus fight" returned over 342,000 results. However, Google Videos returned 3 million results. These include the results identified in the previous YouTube search plus any videos from other sites that meet the search criteria. This will often lead to duplicate videos that have been posted by news websites and social networks. Google can filter these results by duration time, date and time captured, and video source. The top menu of any Google video results page will display these options. A Google search for the term "female street fight", including filters for videos with a short duration that were posted this week from any source, returned over 900 results. These results could either be further filtered with search terms or quickly viewed by still frame to determine relativity to the investigation.

Bing Videos (videos.bing.com)

One feature that makes Bing a favorite site for searching videos is the instant video playback option. When viewing a video search results page, simply hovering the cursor over the video still shot will start the video playback from the beginning of the video. This eliminates the need to navigate to each video page for playback to determine the value of the video. Bing also offers filtering by length and source. The "select view" toolbar at the top of each search result page will allow you to sort the results by either the best match or the most recent. Whether using Google or Bing to locate videos, I recommend turning off the safe search feature. This feature is designed to prohibit some videos with adult content from displaying. With investigations, it is often these types of videos that are wanted.

Real World Application: While investigating an aggravated battery, a local police department searched on Google Videos for the terms "fight" and the city of occurrence. The first result was a YouTube video of the fight taken by an onlooker within five feet of the event. This footage, captured with a cellular phone, also displayed all of the witnesses to the fight. None of the subjects wanted to cooperate, but they were not necessary thanks to the video.

Facebook Videos

Facebook hosts the videos that users embed into their Facebook profiles. If you have already located a target's Facebook page, scrolling through the wall posts will likely present any videos that have been uploaded to the account. If you do not have a specific target and want to search these videos by keywords, this is possible through search engines. Google Videos will include

Facebook videos in its video searches. Including the term Facebook in the search will place an emphasis on videos located on Facebook pages. This will not present every Facebook video fitting the criteria of the search. Using the site operator as discussed in Chapter Three, you can specify a search to include only videos found on Facebook profiles. All Facebook profiles are hosted on the main Facebook domain under a subcategory of video. For example, the address of a video on Facebook would look like this:

http://www.facebook.com/video/video.php?v=1537495389593

The numbers at the end are associated with one unique video. A custom search on Google for a specific video may look like this:

site:facebook.com/video "school bus fight"

This search would attempt to locate any videos stored on the Facebook video servers that included the terms "school bus fight". Most of the videos located should link to the Facebook video page and play without being logged into a Facebook account. After you locate your video of interest, downloading the video can appear difficult. While Facebook does not offer a native download option, you can use their API view to easily archive any content. In the previous Facebook reverse video search discussion, we used the following video page as an example.

https://www.facebook.com/billytalent/videos/10153157582551992/

10153157582551992 is the video ID for this media. Entering the following address into a web browser will reveal a text-only view of the Facebook data for this video.

https://www.facebook.com/video/video_data/?video_id=10153157582551992

At the bottom of this page, you should see a line of code similar to the following.

Hd_source: "https://video-lax3-1.xx.fbcdn.net/v/t43.1792-2/10575621_10153157584221992_398195468_n.mp4?efg=eyJybHIiOjIxNDUsInJsYSI6MjMxNCwidmVuY29kZV90YWciOiJsZWdhY3lfaGQifQ%3D%3D&rl="

The entire URL, beginning with https and ending with .mp4 in this example, can be copied into a web browser address field. The result will be the full screen video playing natively without being embedded into a Facebook page. Saving this page within your browser's file menu will archive the actual video in MP4 format. This presents the best possible quality available without suffering video loss through screen captures or various online download tools. This method provides the purest copy of the target video possible.

World Star Hip Hop (worldstarhiphop.com/videos)

This video sharing website has captured a lot of attention from both the media and law enforcement. The site is infamous for possessing videos of violent fights, sexual activity, and hip-hop music. A search with the term "fight" produced over 400 recent amateur videos, mostly taken with a cellular telephone, depicting recent and brutal street fights with bloody endings. The site averages several million unique visitors every day. Without a doubt, some of those visitors are law enforcement looking to solve cases.

Internet Archive (archive.org/details/opensource_movies)

The premise of this site is to permanently store open source movies, which can include commercial and amateur releases. The search option at the beginning of every page allows for a specific section of the site to be searched. Selecting "community video" will provide the best results for amateur video. A large number of anti-government and anti-American videos are present and ready for immediate download. Unlike YouTube, this site does not make it easy to identify the user that uploaded the videos. Furthermore, it does not link to other videos uploaded by the same user. To do this, you will need to look for some very specific text data. As an example, consider that Internet Archive user Enver_Awlaki is your target. His video profile is located at http://www.archive.org/details/Enver_Awlaki. One of his video pages is stored at the address of https://archive.org/details/Awlaki_to_americans.

Below the large video frame in the center of the page are several options on the lower right. These allow you to specify video files with different file types. Below these is a link titled "Show All". Clicking the link provides a text view of the actual files associated with the video as follows.

Enver_Awlaki.thumbs/	26-Feb-2011 09:08	-
Enver_Awlaki_archive.torrent	26-Jun-2016 22:58	33.4K
Enver_Awlaki_avi.avi	26-Feb-2011 00:05	417.1M
Enver_Awlaki_avi.gif	26-Feb-2011 09:13	308.9K
Enver_Awlaki_avi.ogv	26-Feb-2011 10:46	186.6M
Enver_Awlaki_avi_512kb.mp4	26-Feb-2011 09:52	202.4M
Enver_Awlaki_files.xml	26-Jun-2016 22:58	36.6K
Enver_Awlaki_meta.xml	26-Jun-2016 22:58	686.0B
Enver_Awlaki_wmv.gif	26-Feb-2011 09:04	312.1K
Enver_Awlaki_wmv.ogv	26-Feb-2011 10:19	190.8M
Enver_Awlaki_wmv.wmv	26-Feb-2011 03:18	374.0M
Enver_Awlaki_wmv_512kb.mp4	26-Feb-2011 09:32	202.6M

The eighth link on this list forwards to the metadata associated with the video. This data includes the title, description, creator, email address used to upload, and the date of upload as follows.

```
<mediatype>movies</mediatype>
<collection>opensource_movies</collection>
<title>Awlaki_to_americans</title>
<description>UmmaNews</description>
<subject>UmmaNews</subject>
<creator/>
<identifier>Awlaki_to_americans</identifier>
<uploader>ibnumar@islamumma.com</uploader>
<addeddate>2012-03-31 22:47:36</addeddate>
<publicdate>2012-04-01 00:09:10</publicdate>
```

This view quickly identifies the email address of ibnumar@islamumma.com as the verified uploader of the video content. It also displays the exact date and time of upload and publication. In this example, notice that the author waited over an hour to publish the content. Since 2016, I have seen the Internet Archive become a very popular place to store video, especially from international subjects that may be blocked from traditional American services such as YouTube.

TV News Archive (archive.org/details/tv)

At the time of this writing, the TV News Archive, another part of archive.org, had collected 1,452,000 television news broadcast videos from 2009-2018. Furthermore, it extracts the closed captioning text from each video and provides a search option for this data. This allows you to search for any words verbally stated during these broadcasts in order to quickly locate videos of interest. A search of the term "Bazzell" resulted in 40 videos that mentioned someone with my last name within the broadcast. Selecting any result will play the video and all text from the closed captioning. The menu on the left will allow filtering by show title, station, date, language, and topic. I have found this resource valuable when vetting a potential new hire for a company.

Video Closed Captions (downsub.com)

YouTube and other providers attempt to provide captioning subtitles for as many videos as possible. This automated process transcribes any spoken dialogue within the audio of the video file and documents the words to text. To see this text while watching a video, click on the closed captioning icon (cc) in the lower left area of the video box. When the icon changes to a red color, the subtitles will display. These subtitles are contained within a small text file associated with the video. It also includes time stamps that identify the frame in which each piece of text is spoken. YouTube does not provide a way to obtain this text file, but Downsub does. Copy an entire URL of any YouTube video with closed captioning. Paste this link into this website and execute the process. This will display download links for the captioning inside the video. Links for each language will download text files with an .srt file extension. These automated captions are not usually completely accurate. Slang and mumbled speech may not transcribe properly. Any time that you collect and submit a YouTube video as part of a report, I recommend obtaining this caption file as well. Even though the actual text may not be accurate, it can help during official proceedings with identifying a specific portion of a video.

Vine (vine.com)

Vine was a video sharing service that limited all videos to six seconds in length. It was very popular and many people used it in conjunction with Twitter and Instagram. Vine stopped accepting new videos, but the archives are still available. While Vine offers a reliable search option for videos or users, you should also use Google or Bing. Use the "Site" operator on either service to search for information of interest. The most common search will be for a real name or user name. The following example would query any Vine videos posted by "logan paul".

site:vine.co "logan paul"

You can also search within the text message included with each video. The following search would locate any videos on Vine that have both of the words "fight" and "Nashville".

site:vine.co "fight" "nashville"

Live Video Streams

If you are investigating any live event that is currently occurring, live streaming video sites can be a treasure of useful intelligence. These services offer the ability for a person to turn a cell phone camera into an immediate video streaming device capable of broadcasting to millions. The common setup is for a user to download a host service's application to a smartphone. Launching the application will turn on the video camera of the phone and the video stream is transmitted to the host via the cellular data connection or Wi-Fi. The host then immediately broadcasts this live stream on their website for many simultaneous viewers to see. An average delay time of five seconds is common. There are now several companies that provide this free service. The following are listed in my order of preference for investigative needs. Each site has a search option to enter the keywords that describe the live event you want to watch. You may also see Twitter links to these services while monitoring targets.

UStream (ustream.tv)
LiveStream (livestream.com)
Veetle (veetle.com)
LiveU (liveu.tv)
Bambuser (bambuser.com)
YouNow (younow.com)
VaughnLive (vaughnlive.tv)

Real World Application: During several large events, I have used UStream to capture the majority of my intelligence. In one investigation, I was assigned the task of monitoring social networks during a large protest that had quickly become violent to both officers and civilians. While Twitter and Facebook occasionally offered interesting information, UStream provided immediate vital details that made a huge impact on the overall response of law enforcement, fire, and EMS. The live video streams helped me identify new trouble starting up, victims of violence that needed medical attention, and fires set by arsonists that required an immediate response.

Periscope (pscp.tv)

Periscope is a live video streaming app for iOS and Android acquired in January 2015 by Twitter before the product had been publicly launched. Its use requires a mobile device and there is no web-based official search or player for Periscope streams. Your best option is to search within Twitter. After conducting a keyword search, click the "Broadcasts" option in the top menu. This will present several Periscope video streams. Many will be archived from previous streams, but any live content will be at the top of your results. Once you have identified the video of interest, you should open the media within an official Periscope page. Hover over the video and click on the small Periscope icon in the lower right. It appears similar to a marker on an online map. This will open a page similar to the following.

https://www.pscp.tv/w/1djGXMpDewzJZ

The last portion of this URL is the unique video ID of the target video. This should appear similar to previous instruction on YouTube and Facebook videos. The first search we should conduct in relation to this ID is a Periscope API query. The following URL presents the server data available about the video ID in this example. The Twitter Tools page mentioned in Chapter Five possesses an option to translate an ID to the direct output page.

https://api.periscope.tv/api/v2/getBroadcastPublic?token=1djGXMpDewzJZ

Within the data from this query, we can see the following. Note that this is only a partial list of details, and focused only on the data most valuable to us. The commentary in parentheses explains the purpose of this data.

created_at: 2015-05-13 06:26:59 (Date and time of account creation)
twitter_screen_name: DVATW (Twitter user name)
broadcast created_at: 2017-12-28 19:58:16 (Date and time of video stream creation)
updated_at: 2017-12-28 20:16:10 (Date and time of any profile changes)
friend_chat: false (Target is not using "friend chat")
private_chat: false (Target is not using "private chat")
language: en (Language setting of target's profile)
start: 2017-12-28 20:00:29 (Beginning date and time of video stream broadcast)
has_location: true (User has allowed location enabling)
city: houston (User provided city)
country: USA (User provided country)
country_state: tx (User provided state)
ip_lat: 38.4 (GPS coordinates of IP address in use*)
ip_lng: -90.5 (GPS coordinates of IP address in use*)
width: 320,height:568 (Size of the video stream)
camera_rotation: 270 (Identifies how the phone is rotated)
broadcast_source: periscope_ios_1.13.6 (Identifies the make of his device and app version)

available_for_replay: true (Displays whether user allows archiving of video)
tweet_id: 946470936403775488 (Twitter post announcing video stream)
n_watching: 112 (People currently viewing video)
n_watched: 1140 (Total people that viewed any portion of video)

Note that the GPS coordinates are never meant to be an exact location. These are generic numbers determined from the IP address of the target. These will usually identify the city of the broadcast, but never an accurate address. These settings can be disabled by the user. Worse, use of a VPN could generate unreliable location information. I have had limited success with the following third-party Periscope search services, but nothing can compare to a direct search within Twitter or on the official Periscope page.

Perisearch: perisearch.net
Xxplore: getxplore.com
On Periscope: onperiscope.com

Downloading Videos

In Chapter Two, I mentioned the video download option YouTube-DL, which is pre-configured with the Buscador Virtual Machine. I believe this is the best option for downloading online videos, and it works for both live streams and pre-recorded archives. Simply copy the URL of a video or stream and paste it into the tool. If you are looking for an online solution that does not require any software configuration, I have had success with the following resources.

General Video Download: keepvid.com / clipconverter.cc
YouTube Video Download: videograbby.com / y2mate.com
Facebook Video Download: fbdown.net / download-fb-video.com
Twitter Video Download: downloadtwittervideo.com / twdown.net
Instagram Video Download: downloadvideosfrom.com / w3toys.com
Periscope Video Download: periscopevideodownloader.com / downloadperiscopevideos.com
Vine Video Download: vinevideodownload.com / vinedownloader.com

CHAPTER SIXTEEN
DOMAIN NAMES

A specific web page may quickly become the focus of your investigation. Websites, also known as domains, are the main sites at specific addresses. For example, the website that hosts my blog, www.inteltechniques.com/wp is on the domain of inteltechniques.com. The "www" or anything after the ".com" is not part of the domain. These addresses should always be searched for additional information. If your target has a blog at a custom domain, such as privacy-training.com, the content of the site should obviously be examined. However, digging deeper into the domain registration and associated connections can reveal even more information. Every time that I encounter a domain involved in my investigation, I conduct the following types of research, and usually in the order specified. This chapter will explain techniques for each process.

Current Domain Registration
IP/DNS Configurations
Historical Domain Registration
Live & Historic Visual Depictions
Website Analytics Associations
Server and Content Details
Subdomain Locations
Robots.txt Information
Search Engine Marketing & Optimization
Replication of Content

Domain Registration

Every website requires information about the registrant, administrative contact, and technical contact associated with the domain. These can be three unique individuals or the same person for all. The contact information includes a full name, business name, physical address, telephone number, and email address. These details are provided by the registrar of the domain name to the service where the name was purchased. This service then provides these details to Internet Corporation for Assigned Names and Numbers (ICANN). From there, the information is publicly available and obtained by hundreds of online resources. While ICANN declares that the provided information is accurate, this is rarely enforced. While most businesses supply appropriate contacts, many criminals do not. While we must consider searching this publicly available information, often referred to as Whois details, we will also need to dig deeper into domain analysis in order to obtain relevant results. First, we will focus on the easy queries.

Whois queries (pronounced Who Is) are very simple searches, but are not all equal. While this data is public, it could change often. Some Whois search sites display the bare bones details while others provide enhanced information. There are dozens of options from which to choose, and I will explain those that I have found useful. After the demonstrations, I present my own custom online tool that automates many processes. For every example, I will use a target domain of phonelosers.org, the website of the Phone Losers of America, a telephone hacking and prank calling organization. Assume that this website is the focus of your investigation and you want to

retrieve as much information as possible about the site, the owner, and the provider of the content. For the standard Whois search, as well as many other options, I prefer ViewDNS.info.

ViewDNS Whois (viewdns.info/whois)

This service provides numerous online searches related to domain and IP address lookups. Their main page (viewdns.info) provides an all-in-one toolbox, but the above website connects you directly to their Whois search. Entering phonelosers.org here presents the following information.

Updated Date: 2017-08-17T17:58:11Z
Creation Date: 1997-08-13T04:00:00Z
Registry Expiry Date: 2018-08-12T04:00:00Z
Registrant Name: Brad Carter
Registrant Organization: Phone losers of America
Registrant Street: PO Box 465
Registrant City: Albany
Registrant State/Province: Oregon
Registrant Postal Code: 97321
Registrant Country: US
Registrant Phone: +1.8144225309
Registrant Email: brad@notla.com

The administrative and technical contacts were identical to the registrant shown above. This data identifies Brad Carter as the owner of the site, it was created in 1997, expires in August of 2018, and he has a PO Box in Albany, OR. A telephone number and email address can be searched through the methods explained in previous chapters. This is a great start, if the provided details are accurate. I have found that ViewDNS will occasionally block my connection if I am connected to a VPN. Alternative whois research tools are whois.net and who.is.

Many domain owners have started using private registration services in order to protect their privacy. These services provide their own data within the whois search results, and only these companies know the true registrant. While a court order can usually penetrate this anonymity, I will discuss public resources to help in these situations. While we are discussing ViewDNS, you should be aware of the additional search options available from the main website.

ViewDNS Reverse IP (viewdns.info/reverseip)

Next, you should translate the domain name into the IP address of the website. ViewDNS will do this, and display additional domains hosted on the same server. This service identified the IP address of phonelosers.org to be 104.28.10.123 and stated the web server hosted 134 additional domains. These included domains from websites all over the world without a common theme. This indicates that he uses a shared server, which is very common. If I would have seen only a few domains on the server, that may indicate he is also associated with those specific domains.

ViewDNS Port Scanner (viewdns.info/portscan)

This online port scanner looks for common ports that may be open. An open port indicates that a service is running on the web server that may allow public connection. A search of phonelosers.org revealed that ports 21, 80, and 443 are open to outside connections. Port 80 is for web pages and port 443 is for secure web pages. These are open on practically every website. However, port 21 is interesting. ViewDNS identifies this as a port used for FTP servers, as was discussed in Chapter Three. This indicates that the website hosts an FTP server and connecting to ftp.phonelosers.org could reveal interesting information.

ViewDNS IP History (viewdns.info/iphistory)

This tool translates a domain name to IP address and identifies previous IP addresses used by that domain. A search of phonelosers.org reveals the following details. The first column is the IP address previously associated with the domain, the second column identifies the current user and company associated with that IP address, and the last column displays the date these details were collected by ViewDNS.

104.28.10.123	Reserved	CloudFlare, Inc.	2016-01-24
104.28.11.123	Reserved	CloudFlare, Inc.	2016-01-24
104.28.10.123	Reserved	CloudFlare, Inc.	2016-01-24
208.97.152.79	Brea - United States	New Dream Network, LLC	2015-08-14
162.213.253.190	San Francisco - United States	Namecheap, Inc.	2015-01-15
104.28.10.123	Reserved	CloudFlare, Inc.	2015-01-11
104.28.11.123	Reserved	CloudFlare, Inc.	2014-10-29
104.28.10.123	Reserved	CloudFlare, Inc.	2014-10-17
192.185.46.66	Chelmsford - United States	WEBSITEWELCOME.COM	2014-08-08
74.208.175.23	Wayne - United States	1&1 Internet Inc.	2014-07-04
74.208.211.36	Alliance - United States	1&1 Internet Inc.	2011-05-02

ViewDNS TraceRoute (viewdns.info/traceroute)

This tool identifies the path that ViewDNS took from their servers to the target domain name. This can identify IP addresses of servers that were contacted while you tried to establish communication with the target's website. These will occasionally identify associated networks, routers, and servers. Our target displayed the following results. The IP addresses can be later searched for further details. The numbers after the IP addresses indicate the number of milliseconds that each "hop" took.

```
traceroute to phonelosers.org (104.28.10.123), 30 hops max, 60 byte packets
1 obfuscated.internal.network.com (0.0.0.0) 0.000 ms 0.000 ms 0.000 ms
2 obfuscated.internal.network.com (0.0.0.0) 1.000 ms 1.000 ms 1.000 ms
3 * * *
4 66.231.179.114 (66.231.179.114) 8.246 ms 7.446 ms 8.263 ms
5 66.231.179.86 (66.231.179.86) 7.057 ms 7.085 ms 7.108 ms
6 69.169.95.198 (69.169.95.198) 164.512 ms 155.652 ms 155.599 ms
7 xe-0-6-0-1.r06.asbnva02.us.bb.gin.ntt.net (128.242.179.169) 11.003 ms 8.818 ms 8.831 ms
8 xe-0-6-0-26.r06.asbnva02.us.ce.gin.ntt.net (165.254.106.30) 8.062 ms 8.114 ms 7.247 ms
9 104.28.10.123 (104.28.10.123) 7.377 ms 7.875 ms 7.901 ms
```

Whoisology (whoisology.com)

This service appeared in 2014 and becomes more powerful every month. It is much more than a Whois lookup tool. Technically, it is a *reverse* Whois lookup. However, that description is not powerful enough to convey the search options available within this website. The home page of Whoisology presents a single search field requesting a domain or email address. Entering either of these will display associated websites and the publicly available Whois data. This is where the features begin. The first basic feature that you see is the display of standard Whois data that will identify the registered administrative contact, registrant contact, technical contact, and billing contact. These will often be the same individual for most personal websites. The advanced feature within this content is the ability to immediately search for additional domains associated within any field of this data. As an example, a search for the domain of phonelosers.org reveals the following data.

Name	Brad Carter (88)
Email	brad@notla.com (7)
Street	PO Box 465 (1,091)
City	Albany (42,428)
Region	Oregon (492,506)
Zip / Post	97321 (3,080)
Phone	8144225309 (4)

The name, address, and other data can be found on any whois search website. However, the numbers in parentheses identify the number of additional domains that match those criteria. In this example, there are a total of 88 domains registered to Brad Carter, and seven domains registered to the email address of brad@notla.com. Clicking on any of these pieces of data will launch a new page with all of the matching domain information. As an example, clicking on brad@notla.com will display the 7 domain names associated with his email address. Clicking 8144225309 will display the 4 domain names associated with his telephone number. One of these is a new domain that is not directly associated with him. However, since it was registered with the same number, there is now a connection.

This type of cross-reference search has not been found through other services. Another powerful feature of Whoisology is the historic archives. This service constantly scans for updates to domain registrations. When new content is located, it documents the change and allows you to search the previous data. As an example, a search of computercrimeinfo.com reveals the current administrative contact telephone number to be 6184628253. However, a look at the historic records reveals that on October 16, 2012, the same contact number was 6184633505. This can be a great way to identify associated telephone numbers that have since been removed from the records. Whoisology will also provide details from the search of an email address. In my experience, Whoisology will provide a more detailed and accurate response than most other resources. A search of the email address brad@notla.com revealed the following domains associated with that account.

notla.com
phonelosers.com
albanyscavengerhunt.com
bigbeefbueno.com
callsofmassconfusion.com
snowplowshow.com
phonelosers.org

If you ever encounter an investigation surrounding a domain or any business that possesses a website, I highly encourage you to conduct research through Whoisology. They also offer access through their API at a cost. The individual queries through their website are free.

DNS Trails (dnstrails.com)

This service provides another cross-reference view into the registration data associated with domains. You can easily click on the details, such as the owner name, address, telephone number, or email address, and quickly identify other domains connected to the data. One advantage that I have found with DNS Trails is that it appears to do a better job at connecting organization names with additional domains. In our example, it discovered three additional domains associated with phonelosers.org including cactiradio.com.

Domain History (domainhistory.net)

Domain History is similar to Whoisology and DNS Trails. However, it does not offer many options for cross-reference search of data fields. It does offer a historical view of the Whois registration data as well as related domains based on email address. A search of the domain notla.com revealed the standard Whois data, an associated email address of brad@notla.com, and two additional domains associated with that email address. There were over a dozen historical records of this domain's registration details. Most of them were very recent and identified redundant information. However, one historical record was from six months prior and identified a previous domain registrar. An additional service that offers historical views of domain registrations is domaintools.com, but you must pay for premium access to see all details.

Who Is Hosting This (whoishostingthis.com)

For a quick look at where a website is hosted, try WhoIsHostingThis. Results are minimal and to the point. A search of phonelosers.org revealed CloudFlare as the host and the IP address of the server. This is not a complete search for all information, but this will identify only the key pieces of information without confusion. Law enforcement often serves subpoenas to get information from website hosts. This site offers a start to find the company hosting the content. Once an investigator knows where a website is hosted, then a court order can be sent for more details.

WhoIsMind (whoismind.com)

"Whois" websites are plentiful and rarely unique. They all query user registration data for domain names and provide any hosting details available. WhoIsMind offers additional search options not present on other websites. The standard search page allows for a query of a domain name, IP address, or email address. The domain and IP address searches present identical information as other services. However, an email address search presents any domain names historically associated with the address. This unique option allows you to locate websites that were once registered by your target, before this type of search became routine. A search of brad@notla.com identified a previous registration of the website callsofmassconfusion.com. This service, combined with the email assumptions search described earlier, can lead to the discovery of websites associated with your target's user name.

Visual Depictions

While it may seem obvious, you should document the current visual representation of your target website. This may be a simple screen capture, but you may have advanced needs. When I am investigating a domain, I have three goals for visual documentation. I want documentation of the current site, any historical versions archived online, and then future monitoring for any changes. Chapters One and Two presented numerous methods for capturing live web pages, and Chapter Three discussed several techniques for locating archives online such as the Wayback Machine and Google Cache. You should now consider checking the next resource.

Screenshots (screenshots.com)

While search engines provide the most recent cached version of a web page, and the Wayback Machine creates backups of websites, Screenshots takes a unique route. This service visits domains regularly and creates a screen capture image. These are all stored and available to the public. In our example, Screenshots offered 34 historic views of phonelosers.org ranging from 7/29/2004 through 11/19/2017. These third-party captures can prove valuable in court as unbiased references to a target domain. This resource services as a great compliment to traditional cached and archived websites. Now that we have an understanding of historic data, this leaves us with the requirement to monitor your target website for future changes.

Follow That Page (followthatpage.com)

Once you locate a website of interest, it can be time consuming to continually visit the site looking for any changes. With large sites, it is easy to miss the changes due to an enormous amount of content to analyze. This is when sites like Follow That Page come in handy. Enter the address of the target page of interest as well as an email address where you can be reached. This service will monitor the page and send you an email if anything changes. Anything highlighted is either new or modified content. Anything that has been stricken through indicates deleted text. Parents are encouraged to set up and use these services to monitor their child's websites. It does not work well on some social networks such as Facebook, but can handle a public Twitter page fine.

Visual Ping (visualping.io)

If you found the service provided by Follow That Page helpful, but you are seeking more robust options, you should consider Visual Ping. This modern Swiss website allows you to select a target domain for monitoring. Visual Ping will generate a current snapshot of the site and you can choose the level of monitoring. I recommend hourly monitoring and notification of any "tiny change". It will now check the domain hourly and email you if anything changes. If you are watching a website that contains advertisements or any dynamic data that changes often, you can select to avoid that portion of the page. Figure 16.01 displays the monitoring option for phonelosers.org. In this example, I positioned the selection box around the blog content of the main page. I also chose the hourly inspection and Tiny Change option. If anything changes within this selected area, I will receive an email announcing the difference.

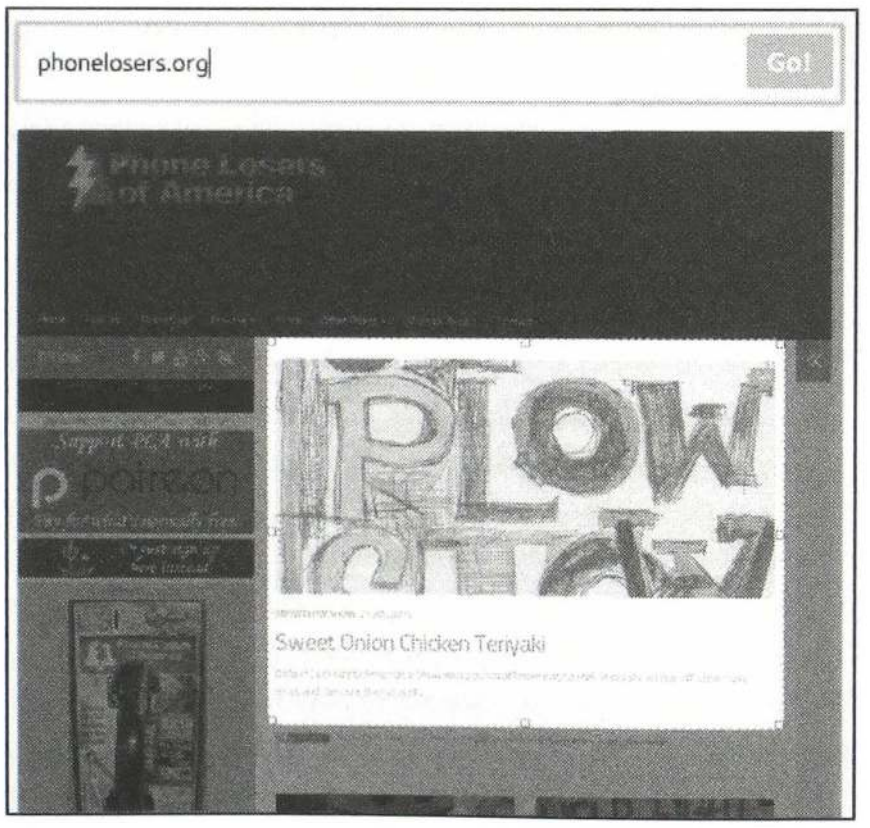


Figure 16.01: A portion of a web page monitored by Visual Ping for changes.

Reverse Domain Analytics

Domain analytics are commonly installed on websites in order to track usage information. This data often identifies the city and state from where a visitor is; details about the web browser the person is using; and keywords that were searched to find the site. Only the owner of the website can view this analytic data. Analytics search services determine the specific number assigned to the analytics of a website. If the owner of this website uses analytics to monitor other websites, the analytic number will probably be the same. These services will now conduct a reverse search of this analytic number to find other websites with the same number. In other words, it will search a website and find other websites that the same owner may maintain. Additionally, it will try to identify user specific advertisements stored on one site that are visible on others. It will reverse search this to identify even more websites that are associated with each other. None of this relies on Whois data. A couple of examples should simplify the process.

Spy On Web (spyonweb.com)

Spy On Web is one of many sites that will search a domain name and identify the web server IP address and location. It also conducts a Whois query which will give you registration information of a website. More importantly, it identifies and cross-references website analytic data that it locates on a target domain. A search for the website phonelosers.org reveals a "Google AdSense" ID of pub-3941709854725695. It further identifies five domains that are using the same Google AdSense account for online advertising. This identifies an association between the target website and these new websites. We now know that whoever maintains phonelosers.org places ads on the page. We also know that those same ads and affiliate number is present on five domains. This means that our target likely maintains all of the following domains.

www.notla.com
www.oldpeoplearefunny.com
www.phonelosers.com
www.phonelosers.org
www.signhacker.com

Analyze ID (analyzeid.com)

While Spy On Web is a strong overall analysis tool, there are additional options that should be checked for reverse analytics. Analyze ID performs the same type of query and attempts to locate any other domains that share the same analytics or advertisement user numbers as your target. This will provide new websites related to your target. During a search of phonelosers.org, it identified the Google AdSense ID mentioned previously. It also revealed an Amazon Affiliate ID of phonelosersof-20 and an Amazon product ID of 1452876169. These are likely present because the target sells books through his websites and receives compensation from Amazon through customer purchases. These new pieces of information can be very valuable. Clicking the Amazon Affiliate ID presents the five domains that possess that same code. The Amazon product ID displays the seven websites that also advertise a specific product. When you encounter this type of data, consider the following Google search. It identifies the actual product

that the target is selling within embedded ads. This search reveals that the product is the target's self-published book titled Phone Losers of America.

amazon product 1452876169

Analyze ID also identified a Clickbank affiliate ID that was associated with five additional domains, one of which was unique to any other results. Spy On Web and Analyze ID have provided us several new pieces of information about our target. If you ever see an ID that starts with "UA-", this is likely an identifier for Google to monitor viewers of the website. Searching that number within these tools will also identify related websites. We should visit each website and analyze the content for any relevant evidence. After, we should consider other resources.

PubDB (pub-db.com)

Another free domain analysis service is PubDB. It is not as robust as others mentioned; however, it occasionally identifies information that was not available with other services. It identified our target's Google AdSense affiliate number within a website and cross-referenced that data to three other related domains.

Domain Crawler (domaincrawler.com)

The results on this service will likely be redundant to the previously mentioned websites. However, the data here could be used to validate the information that you have already collected. Domain Crawler will attempt to identify any analytics or advertisement codes and will allow you to cross-reference these to other websites. A search for computercrimeinfo.com only identified three related websites. The results were accurate.

Nerdy Data (search.nerdydata.com)

Nerdy Data is a search engine that indexes the source code of websites. If you have located a Google Analytics, AdSense ID, or Amazon ID of a website using the previous methods, you should consider searching this number through Nerdy Data. A search of our target's Google AdSense number revealed five domains that possess the same data. The search of the Amazon number revealed three domains. If this service presents more results than you can manage, consider using their free file download option to generate a csv spreadsheet.

Real World Application: While investigating an "anonymous" website that displayed photo evidence of a reported felony, I discovered that the registration information was intentionally inaccurate. A search of the website on these services identified a Google analytics number and an additional website that possessed the same number. That additional website was the personal blog of the suspect. An arrest was made the same day.

Built With (builtwith.com)

A quick analysis of a target website may identify the technologies used to build and maintain it. Many pages that are built in an environment such as WordPress or Tumblr often contain obvious evidence of these technologies. If you notice the YouTube logo within an embedded video, you will know that the creator of the site likely has an account within the video service. However, the presence of various services is not always obvious. Built With takes the guesswork out of this important discovery. Entering the domain of phonelosers.org into the Built With search immediately identifies the web server operating system (Linux), email provider (DreamHost), web framework (PHP, WordPress), WordPress plugins, website analytics, video services, mailing list provider, blog environment, and website code functions. While much of this is geek speak that may not add value to your investigation, some of it will assist in additional search options through other networks. Other options for this type of search include **Stats Crop** (statscrop.com) and **URL Scan** (urlscan.io).

Pentest-Tools (pentest-tools.com/reconnaissance/find-subdomains-of-domain)

This unique tool performs several tasks that will attempt to locate hidden pages on a domain. First it performs a DNS zone transfer which will often fail. It will then use a list of numerous common subdomain names and attempt to identify any that are present. If any are located, it will note the IP address assigned to that subdomain and will scan all 254 IP addresses in that range. In other words, it will attempt to identify new areas of a website that may not be visible from within the home page. The following example may help to clarify.

The website at phonelosers.org is a blog that appears to have no further content to be analyzed. Searching for it on Pentest-Tools provides additional intelligence. It identifies the following subdomains present on the web server:

webmail.phonelosers.org
ssh.phonelosers.org
ftp.phonelosers.org
www.phonelosers.org
mail.phonelosers.org

We now know that this domain possesses a webmail server, SSH connection, FTP server, and mail server. This method has helped me locate "hidden" pages which contain several forum messages from users of the site. Previous editions of this book have discussed additional providers for this type of service. Pentest-Tools is the only provider that continues to function. The rest have disappeared.

Robots.txt

Practically every professional website has a robots.txt file at the "root" of the website. This file is not visible from any of the web pages at the site. It is present in order to provide instructions to search engines that crawl the website looking for keywords. These instructions identify files and

folders within the website that should not be indexed by the search engine. Most engines comply with this request, and do not index the areas listed. Locating this file is relatively easy. The easiest way to view the file is to open it through a web browser. Type the website of interest, and include "robots.txt" after the forward slash (/). The file for Reddit can be found at the following address.

http://www.reddit.com/robots.txt

If this technique produces no results, you can conduct a Google or Bing query to identify any files. A search of site:twit.tv "robots.txt" on either search engine identifies robots.txt files from the entire website Twit.tv. Different robots.txt files were found at both twit.tv and inside.twit.tv. The main robots.txt located at twit.tv/robots.txt appears as the following.

```
User-agent: *
Crawl-delay: 10
Sitemap: https://twit.tv/sitemap.xml
```

This was used as an interesting way to show a message to curious web visitors while providing search engines a sitemap for accurate indexing of their site. However, the robots.txt file located at inside.twit.tv/robots.txt provides more lucrative information, as seen below.

```
# Squarespace Standard Robot Exclusion
# Access is disallowed to functional / filtering URLs

User-agent: *
Disallow: /display/admin/
Disallow: /display/Search
Disallow: /display/Login
Disallow: /display/RecoverPassword
Disallow: /login
Disallow: /contributor

Disallow: /blogold/category
Disallow: /blogold/week
Disallow: /blogold/month
Disallow: /blogold/recommend
Disallow: /blogold/author
Disallow: /login
Disallow: /blog/category
Disallow: /blog/week
Disallow: /blog/month
Disallow: /blog/recommend
Disallow: /blog/author
Disallow: /contests/category
Disallow: /contests/week
Disallow: /contests/month
Disallow: /contests/recommend
Disallow: /contests/author
```

The first line indicates that this website was created by Squarespace, and is likely stored on their servers. The rest of the file identifies online folders that include a live blog, previous blog, contest area, and login portal. Much of this content would not be found in a search engine because of the Disallow setting. These Disallow instructions are telling the search engines to avoid scanning the folders login, contributor, and RecoverPassword. It is likely that there is sensitive information in these directories that should not be available on Google or Bing. You can now type these directories after the domain name of your target to identify additional information. Typing the following addresses directly into a browser generates interesting results.

http://inside.twit.tv/blog
http://inside.twit.tv/blogold
http://inside.twit.tv/display/RecoverPassword

Most robots.txt files will not identify a secret area of a website that will display passwords, raunchy photos, or incriminating evidence. Instead, they usually provide insight into which areas of the site are considered sensitive by the owner. If you have a target website and have exhausted every other search method, you should also visit this file. It may direct you toward a new set of queries to find data otherwise ignored by search engines.

Search Engine Marketing Tools

The ultimate goal of most commercial websites is to generate income. These sites exist to bring in new customers, sell products, and provide the public face to a company. This has created a huge community of services that aim to help companies reach customers. Search Engine Optimization (SEO) applies various techniques affecting the visibility of a website or a web page in a search engine's results. In general, the higher ranked on the search results page and more frequently a site appears in the search results list, the more visitors it will receive. Search Engine Marketing (SEM) websites provide details valuable to those responsible for optimizing their own websites. SEM services usually provide overall ranking of a website, its keywords that are often searched, backlinks, and referrals from other websites. SEO specialists use this data to determine potential advertisement relationships and to study their competition. Online investigators can use this to collect important details that are never visible on the target websites. Three individual services will provide easily digestible data on any domain. I will use my own domain for each example in order to compare the data. Only the free versions will be discussed.

SEM Rush (semrush.com)

SEM Rush is usually the most comprehensive of the free options. Entering inteltechniques.com as the target produced the following partial information about the domain.

The majority of the traffic is from the USA, followed by UK, DE, FR
There is no paid search or advertisements on search engines
There are 56 websites that possess links to the target, and 15 are visible

"Facebook Search" led more people to the site than any other search term followed by OSINT
Over 5,000 people visit the site monthly
There are five main online competitors to the target, and the largest is onstrat.com

Similar Web (similarweb.com)

Similar Web provides a similar view with redundant information. However, some of these details usually contradict other services. Much of this data is "guessed" based on many factors. A search of inteltechniques.com produced the following partial information about the domain.

71% of the visitors navigated directly to the domain without a search engine
12% of the traffic was referrals from other websites and search engines
The referrals included my other website (computercrimeinfo.com)
The top destination sites visited after the target domain included Facebook and Pipl
Top search terms included IntelTechniques, OSINT, and Michael Bazzell
3% of the traffic to this site originated from the social networks Facebook, Twitter, and Reddit
Similar websites include onstrat.com and automatingosint.com

Alexa (alexa.com)

This service is considered the standard when citing the global rank of a website. Most of the collected data is targeted toward ranking the overall popularity of a website on the internet. The following details were provided about inteltechniques.com

It is ranked as the 330,033rd most popular site on the internet.
The average visitor clicks on three pages during a visit.
Popular searches used to find the domain include OSINT Links and IntelTechniques.
Facebook, Twitter, and Google referred more traffic than any other source.

All three of these services provided some details about the target domain that were not visible within the content of the page. These analytical pieces of data can be valuable to a researcher. Knowing similar websites can lead you to other potential targets. Viewing sources of traffic to a website can identify where people hear about the target. Global popularity can explain whether a target is geographically tied to a single area. Identifying the searches conducted before reaching the target domain can provide understanding about how people engage with the website. While none of this proves or disproves anything, the intelligence gathered can help give an overall view of the intent of the target. Three additional websites that provide a similar service are listed below.

Search Metrics (suite.searchmetrics.com)
Majestic (majestic.com)
SpyFu (spyfu.com)

Shared Count (sharedcount.com)

This website provides one simple yet unique service. It searches your target domain and identifies its popularity on social networks such as Facebook and Twitter. A search of labnol.org produced the following results.

Facebook Likes: 348
Facebook Shares: 538
Facebook Comments: 148
Facebook Total: 1034
Twitter Tweets: 0
Google+1s: 4202
Pinterest Pinned: 1
LinkedIn Shares: 172
Delicious Bookmarks: 44
StumbleUpon Stumbles: 0

This information would lead me to focus on Google+ and Facebook first. It tells me that several people are talking about the website on these services. I also know now that 44 people have bookmarked the site on Delicious, and I could go track down those people. I have used this tool to successfully identify pedophiles that are interested in a child pornography website and students at a high school that were commenting on a blog that encouraged harassment of a specific student.

Small SEO Tools: Backlinks (smallseotools.com/backlink-checker)

After you have determined the popularity of a website on social networks, you may want to identify any websites that have a link to your target domain. This will often identify associates and people with similar interests of the subject of your investigation. There are several online services that offer a check of any "backlinks" to a specific website. Lately, I have had the best success with the backlink checker at Small SEO Tools. A search of my own website, inteltechniques.com, produces 264 websites that have a link to mine. These results include pages within my own websites that have a link to inteltechniques.com, so this number can be somewhat misleading. Several of the results disclosed websites owned by friends and colleagues that would be of interest if I were your target.

Small SEO Tools: Plagiarism Checker (smallseotools.com/plagiarism-checker)

If you have identified a web page of interest, you should make sure that the content is original. On more than one occasion, I have been contacted by an investigator that had been notified of a violent threat on a person's blog. I was asked to track down the subject before something bad happened. A quick search of the content identified it as lyrics to a song. One of many options for this type of query is the plagiarism checker at Small SEO Tools.

You can use this tool by copying any questionable text from a website and paste it into this free tool. It will analyze the text and display other websites that possess the same words. This service uses Google to identify anything of interest. The benefit of using this tool instead of Google directly is that it will structure several queries based on the supplied content and return variations

of the found text. Clicking the results will open the Google search page that found the text. Another option for this type of search is **Copy Scape** (copyscape.com).

Reddit Domains (reddit.com)

Reddit was discussed in Chapter Seven as a very popular online community. The primary purpose of the service is to share links to online websites, photos, videos, and comments of interest. If your target website has ever been posted on Reddit, you can retrieve a listing of the incidents. This is done through a specific address typed directly into your browser. If your target website was phonelosers.org, you would navigate to the following website.

reddit.com/domain/phonelosers.org/

This example produced 16 Reddit posts mentioning this domain. These could be analyzed to document the discussions and user names related to these posts.

Hunter (hunter.io)

In Chapter Eight, I explained how Hunter could be used to verify email addresses. This tool can also accept a domain name as a search term, and provides any email addresses that have been scraped from public web pages. The free version of this tool will redact a few letters from each address, but the structure should be identifiable.

Visual Site Mapper (visualsitemapper.com)

When researching a domain, I am always looking for a visual representation to give me an idea of how massive the website is. Conducting a "Site" search on Google helps, but you are at the mercy of Google's indexing, which is not always accurate or recent. An alternative to this is to use Visual Site Mapper. This service analyzes the domain in real time, looking for linked pages within that domain. It provides an interactive graph that shows whether a domain has a lot of internal links that you may have missed. Highlighting any page will display the internal pages that connect to the selected page. This helps identify pages that are most "linked" within a domain, and may lead a researcher toward those important pages. Figure 16.02 displays a portion of the map for our previous target. I hovered over a single page, which identifies the URL and highlights any internal pages with links pointing back to it. This visual representation helps me digest the magnitude of a target website.

Virus Total (virustotal.com)

This chapter would not be complete without mentioning Virus Total. While the service is focused on the analysis of malicious software, files, and websites, a search here will also identify historic DNS records, whois data, subdomains, and any suspicious URLs. Queries are straight forward, and all available results are freely visible after each search.

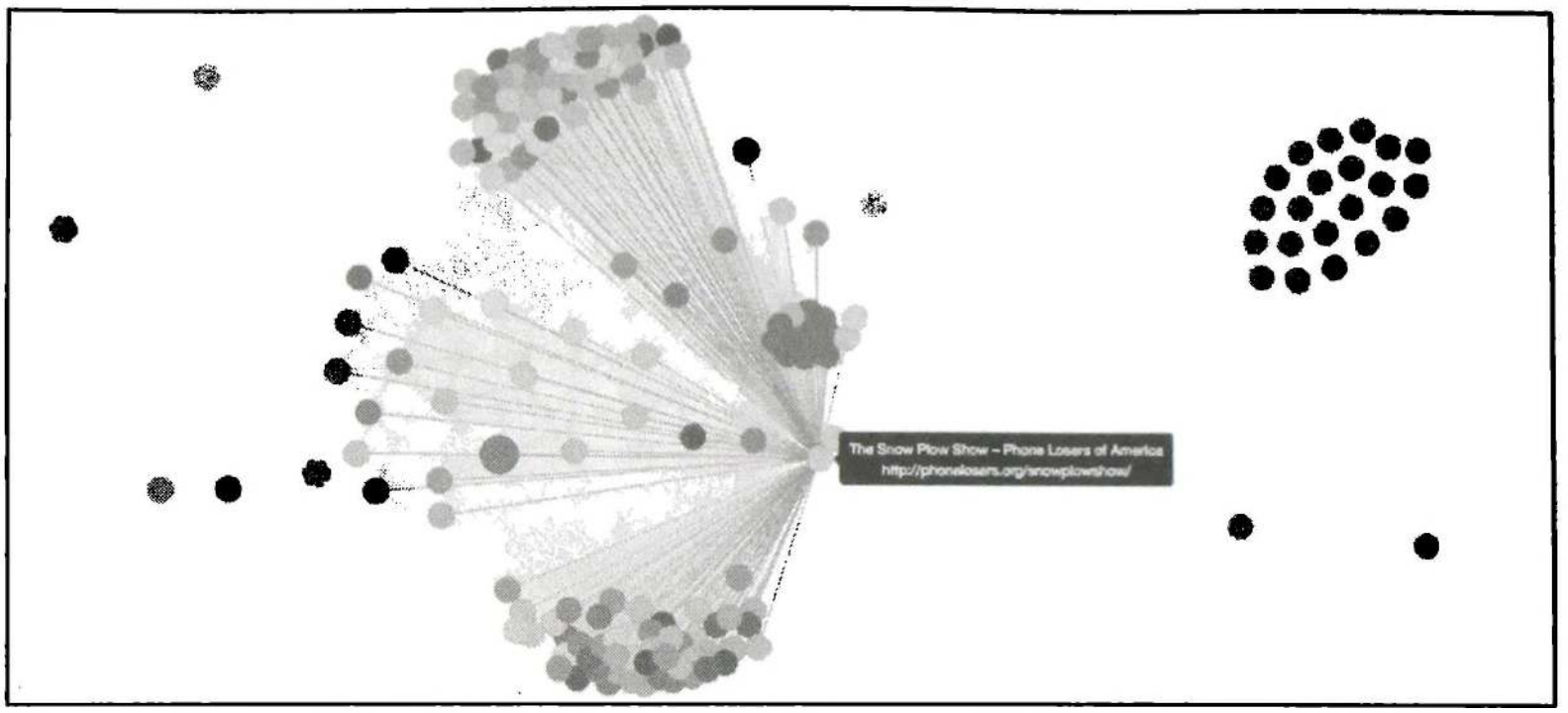


Figure 16.02: A partial view of a website mapping from Visual Site Mapper.

Shortened URLs

Social networking sites such as Twitter have made the popularity of shortened URL services soar. When people post a link to something they want their friends to see, they do not want the link to take up unnecessary space. These services create a new URL, and simply point anyone to the original source when clicked. As an example, I converted a URL to a blog post from this:

inteltechniques.com/wp/2017/12/29/the-complete-privacy-security-podcast-episode-060/

to this: bit.ly/2CluQOK

You have likely seen these during your own investigations, and many people pay them little attention. There is actually a lot of information behind the scenes of these links that can reveal valuable information associated with your investigation. For a demonstration, I created the following shortened links, all of which forward to my home page. After, I will explain how to access the hidden data behind each service.

bitly.com/29A4U1U
tiny.cc/p20scy
goo.gl/Ew9rlh
bit.do/cbvNx

Bitly allows access to metadata by including a "+" after the URL. In our scenario, the direct URL would be bitly.com/29A4U1U+. In this example, the results only identified that 21 people have clicked on my link. However, creating a free account reveals much more detail. After logging in,

I can see any websites that referred the user to the link and extremely generic location data, such as the country of the user. This is a good start.

Tiny.cc adds a "~" to the end of a link to display metadata. In our example, the direct URL would be tiny.cc/p20scy~. The results on this page identify the number of times the URL was clicked, the number of unique visits, the operating systems of those that clicked the link, and the browsers used. This service also displays generic location data, such as the country of the user.

Google gives us the same detail as above. It also uses the "+" at the end, and our direct demo URL would be goo.gl/Ew9rlh+. This demo notified me that 18 people have clicked my link from 7 different countries. They are mostly Windows users with the Chrome browser.

Bit.do provides the most extensive data. They use a "-" after the URL, and our direct demo address would be http://bit.do/cbvNx-. The results identify all of the details listed previously, plus the actual IP addresses of each visit. In this demo, I know the following about those that clicked on my trap. Note that I redacted the IP addresses to respect the privacy of those involved.

User's IP	Country/City	Access Date
173.244.48.111	United States (Los Angeles, California)	2017-12-29 12:58:55
213.133.92.111	Cyprus (Nicosia, Nicosia)	2017-10-31 04:00:19
125.31.39.111	Macau	2017-04-22 03:36:57
198.8.80.111	United States (Los Angeles, California)	2017-01-12 11:27:12
195.235.92.111	Spain (Madrid, Madrid)	2016-11-18 03:33:22

This type of service can be used in many ways. If you are investigating a viral Twitter post with a shortened URL, you may be able to learn more about the popularity and viewers. You could also use this offensively. During covert investigations, you could forward a shortened URL from Bit.do and possibly obtain the IP address being used by the suspect. I will explain more options for this type of use in the next chapter. If you are investigating a shortened URL link that was not mentioned, consider using the catch-all service at **CheckShortURL** (checkshorturl.com).

IntelTechniques Domain Search Tool (inteltechniques.com/intel/OSINT/domain.search.html)

Similar to the previous custom search tools hosted on IntelTechniques.com, I have created a page for easier domain searching. While it does not possess every service discussed here, it can automate queries across the most beneficial options. The first box accepts any domain name, preferably without www, http, or other leaders. Clicking the Populate All button will insert this domain into all of the search options where manual queries can be conducted. The "Submit All" option will open several tabs within your browser that present each query listed on the page. The final section provides easy access to shortened URL metadata. Figure 16.03 displays the current status of the service.

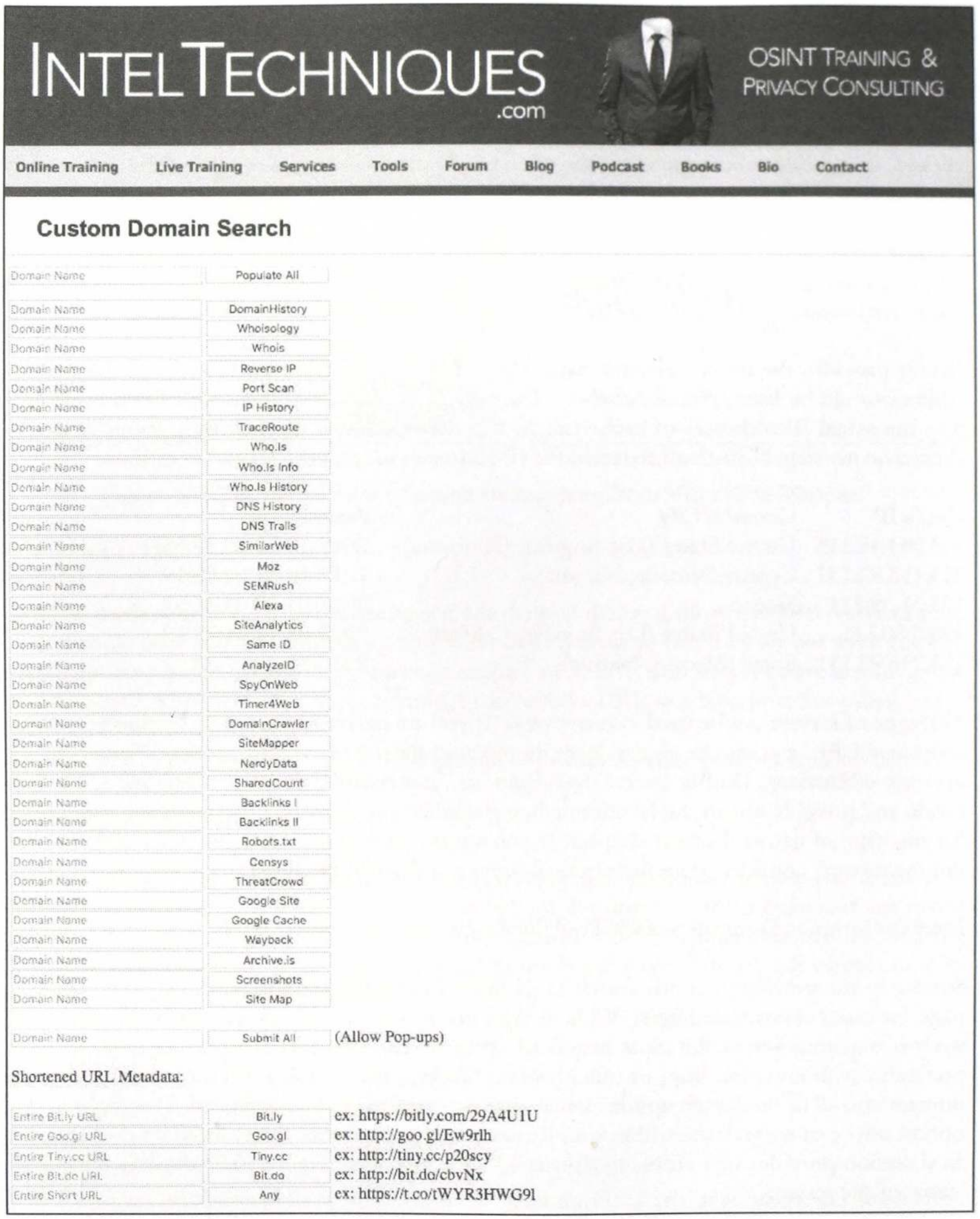


Figure 16.03: The IntelTechniques Custom Domain Search Tool.

CHAPTER SEVENTEEN
IP ADDRESSES

IP addresses are often obtained from an internet investigation, email message, or connection over the internet. When legal process is served to online content providers, a list of IP addresses used to log into the account is usually presented as part of the return of information. Serving legal orders to identify and obtain IP addresses is outside the scope of this book. However, several techniques for collecting a target's IP address using OSINT are explained in this chapter. The previous instruction assumed that you were researching a domain name. These names, associated with websites, simply forward you to a numerical address that actually hosts the content. This is referred to as an Internet Protocol (IP) address.

If you know the IP address of a website, it can be entered into an address field instead of the domain name. As an example, when you enter google.com into a browser, your connection forwards you to 74.125.224.72. Typing this IP address directly into your browser will present the Google landing page. The way that you encounter IP addresses as a target of your research will vary widely. Law enforcement may receive an IP address of an offender after submitting a subpoena to an internet provider. Any online researcher may locate an IP address while researching a domain with the previous methods. While only one website can be on a domain, multiple domains can be hosted on one IP address. The following resources represent only a fraction of available utilities. Note that many of the domain name resources mentioned in the previous chapters also allow for query by an IP address.

IPLocation (iplocation.net)

IPLocation offers unlimited free IP address searches, and queries five unique services within the same search results. The results are the most comprehensive I have seen for a free website. While GPS coordinates of an IP address are available, this most often returns to the provider of the internet service. This usually does not identify the exact location of where the IP address is being used. The country, region, and city information should be accurate. If an organization name is presented in the results, this indicates that the address returns to the identified company. The exception here is when an internet service provider is identified. This only indicates that the IP address belongs to the specified provider. Most results translate an IP address into information including business name, general location, and internet service provider. This can be used to determine if the IP address that a target is using belongs to a business providing free wireless internet. If you see "Starbucks", "Barnes & Noble", or other popular internet cafes listed in the results, this can be important intelligence about the target. This can also quickly confirm if a target IP address is associated with a VPN service. Alternative websites include **IP Fingerprints** (ipfingerprints.com) and **IP2Location** (ip2location.com).

Bing IP (bing.com)

Once you have identified an IP address of your target, you can search for websites hosted on that IP address. A specific search on Bing will present any other websites on that server. If your target is stored with a large host such as GoDaddy, there will not be much intelligence provided. It will only list websites that share a server, but are not necessarily associated with each other. If the user is hosting the website on an individual web server, this search will display all other websites that the user hosts. This search only works on Bing and must have "ip:" before the IP address. An example of a proper search on Bing would look like ip:54.208.51.71. The results of this search identify every local website hosted by a specific local website design company.

ViewDNS Reverse IP (viewdns.info/reverseip)

This page was used previously to translate a domain name into an IP address. It will also display additional domains hosted on an individual IP address. This service identified 134 domains hosted on 104.28.10.123. These included domains from websites all over the world without a common theme. This indicates that he uses a shared server, which is very common. If I would have seen only a few domains on the server, that may indicate that he is also associated with those specific domains.

ViewDNS IP Location (viewdns.info/iplocation)
This utility cross-references an IP address with publicly available location data connected to the server hosting any domains associated with the IP address. A search of 54.208.51.71 revealed the following information.

City:	Ashburn
Zip Code:	20147
Region Code:	VA
Region Name:	Virginia
Country Code:	US
Country Name:	United States

ViewDNS Port Scan (viewdns.info/portscan)

This online port scanner looks for common ports that may be open. An open port indicates that a service is running on the web server that may allow public connection. A search of 54.208.51.71 revealed that ports 21, 53, 80, and 443 are open to outside connections. Port 21 is for FTP connections, 53 is for DNS settings, 80 is for web pages, and port 443 is for secure web pages.

ViewDNS IP Whois (viewdns.info/whois)

This service was used earlier to display registration information about an individual domain. Entering an IP address will attempt to identify details about any domain registrations associated

with the address. A search of 54.208.51.71 revealed it to belong to Amazon and provided the public registration details.

ViewDNS IP TraceRoute (viewdns.info/traceroute)

This tool identifies the path that ViewDNS took from their servers to the target IP address. This can identify IP addresses of servers that were contacted while you tried to establish communication with the target's address. These will occasionally identify associated networks, routers, and servers. Additional IP addresses can be later searched for further details. The numbers after the IP addresses indicate the number of milliseconds that each "hop" took.

That's Them (thatsthem.com/ip)

The previous resources rely on conventional IP address data, which is sourced from various registration documentation and scanning of servers. Very little information is sensitive or personal in nature. That's Them enters into an environment that is a lot more invasive. This service, mentioned previously during person, email, and telephone search, collects marketing data from many sources to populate its database. This often includes IP address information. These details could have been obtained during an online purchase or website registration. Regardless of the source, the results can be quite beneficial. At the time of this writing, I searched an IP address associated with a business email that I received. The result identified a person's name, home address, company, email address, and age range. All appeared accurate. This tool will work best when searching static business IP addresses, and not traditional home addresses that can change often. While I get no results much more often than positive results, this resource should be in everyone's arsenal.

I Know What You Download (iknowwhatyoudownload.com)

While discussing invasive websites, this resource might be the most personal of all. This service monitors online torrents (ways to download large files which often violate copyright laws) and discloses the files associated with any collected IP addresses. I searched the previous IP address collected from an associate, and received an immediate hit. Figure 17.01 displays the result. It identifies that the target IP address was downloading two specific movies on December 28, 2017 at 9:53 pm. Clicking on the movie title presents every IP address captured that also downloaded the same file. Again, this will work best with IP addresses that rarely change, such as a business, organization, or public Wi-Fi network. I have used this to determine the files being downloaded from the network with which I was currently connected. On one occasion, this revealed an employee that was downloading enormous amounts of pornography on his employee's network. He should have used a VPN, which would have masked his online activity from me. In order to see the power of this type of service, try searching a known VPN address such as an address provided by Private Internet Access (PIA) 173.244.48.163. While I know that no one reading this book has ever downloaded pirated content, this should serve as a reminder why VPNs are essential.

Dec 28, 2017 9:53:21 PM	Dec 28, 2017 9:53:21 PM	Movies	Looper
Dec 28, 2017 9:53:19 PM	Dec 28, 2017 9:53:19 PM	Movies	The Beguiled

Figure 17.01: A search result from I Know What You Download.

Exonerator (exonerator.torproject.org)

The Onion Router (Tor) was explained in Chapter Two. It is a network that provides anonymity by issuing IP addresses to users that connect to servers in other countries. If you possess an IP address of your target, but cannot locate any valuable information using the previous techniques, it is possible that the address was part of the Tor network and there is no relevant data to be located. Exonerator is a tool that will verify the usage of an IP address on the Tor network. Provide the IP address and a date of usage, and the service will display whether it was used as a Tor connection. While a date is required, you could provide the current date if your target time frame is unknown. Most IP addresses are either always a part of the Tor network or not connected at all.

Wigle (wigle.net)

Wigle is a crowd sourced database of wireless access points. Users in all areas of the country conduct scans of wireless devices in their area; identify details of each device; and submit this data to Wigle in order to map the found devices on the site. This allows anyone to browse an area for wireless access points or search an address to locate specific devices. Additionally, you can search for either a specific router name or MAC address and locate any matching devices. The results will include links that will display the results on an interactive map. Most of the world has been covered. In order to take advantage of the search features, you will need to register for a free account. Generic or misleading information can be used that does not identify you.

There are many investigative uses for this service. You can identify the wireless access points in the immediate area of a target's home. As an example, a search of the address of a gas station revealed a map of it with the associated routers. In this view, I can identify the router names including possibly sensitive information. It displays wireless router SSID's of AltonBPStore, tankers_network, Big Toe, and others. Clicking View and then Search in the upper left of the page presents a detailed query engine. A search of tankers_network, as identified previously in the map view, displays details of the wireless access point. It has a MAC address of 00:1F:C6:FC:1B:3F, WPA encryption, was first seen in 2011, and operates on channel 11. An investigator could also search by the target's name. This may identify routers that have the target's name within the SSID. A search of "Bazzell" identifies seven access points that probably belong to relatives with my last name. These results identify the router name, MAC address, dates, encryption method, channel, and location of the device. This can easily lead an investigator to the home of a target.

Many internet users will use the same name for their wireless router as they use for their online screen name. Assume that your target's user name was "Hacker21224". A search on Wigle for "Hacker21224" as a router name might produce applicable results. These could identify the router's MAC address, encryption type, and GPS coordinates. A search on Google Maps of the supplied GPS coordinates will immediately identify the home address, a satellite view of the neighborhood, and a street view of the house of the target. All of this intelligence can be obtained from a simple user name. These results would not appear on any standard search engines.

Shodan (shodan.io)

Shodan is a search engine that lets you find specific computers (routers, servers, etc.) using a variety of filters. General search engines, such as Google and Bing, are great for finding websites; however, they do not search for computers or devices. Shodan indexes "banners", which are metadata that a device sends back to a client. This can be information about the server software, what options the service supports, or a welcome message. Devices that are commonly identified through Shodan include servers, routers, online storage devices, surveillance cameras, webcams, and VOIP systems. Network security professionals use this site to identify vulnerabilities on their systems. Criminals use it to illegally access networks and alter devices. We will use it to locate specific systems near a target location. In order to take advantage of Shodan's full search capabilities, you must create a free account. Only a name and email address is required. The following example will identify how to locate live public surveillance cameras based on location. The target for this search is Mount Pleasant, Utah. The following search on Shodan produced 9,684 results.

country:US city:"Mount Pleasant"

There are two flaws with this search. First, you may receive results from other cities named Mount Pleasant. Second, you will likely receive too many results to analyze effectively. A search of "geo:39.55,-111.45" will focus only on the specific GPS location of interest (Lat=39.55, Long=-111.45). There were 238 results for this search. This is much more manageable and all of the results will be devices in the target area. Adding more specific search criteria will filter the results further. A search of "geo:39.55,-111.45 netcam" identified only one device.

The result identifies this device as a "Netcam". It also identifies the internet service provider as "Central Utah Telephone" indicating the user has a DSL connection. To connect to the device, you would click on the IP address identified as 63.78.117.229. Clicking through each of these results may be time consuming. You can add a search term to filter your results. Replicating this search for a GPS location in a large city will produce many results. Clicking the IP address will take you to the page that will connect to each device. You must be careful here. Some devices will require a user name and password for access. You could try "admin" / "admin" or "guest" / "guest", but you may be breaking the law. This could be considered computer intrusion. However, many of the webcam and netcam results will not prompt you for a password and connect you to the device automatically. There is likely no law violation when connecting to a

device that does not prompt you for credentials. Your local laws may prohibit this activity. **Shodan Maps** (maps.shodan.io) allows you to conduct any of these searches based on location alone while **Shodan Images** (images.shodan.io) displays collected webcam captures from open devices. Figure 17.02 displays a home using an automated lighting and climate control system in Missouri located with Shodan Maps. These two options are premium services and require a modest fee. All Shodan features allow input of the following types of information for filtering.

City: Name of the city (ex. city:"San Diego")
Country: 2-letter country code (ex. country:US)
GPS: Latitude and longitude (ex. geo:50.23,20.06)
OS: Operating system (ex. os:Linux)
IP Address: Range (ex. net:18.7.7.0/24)
Keyword: (ex. webcam)



Figure 17.02: A Shodan Maps search result.

ZoomEye (zoomeye.org)

In 2016, ZoomEye surfaced as a popular Shodan clone. While the same search parameters used in Shodan do not work reliably in ZoomEye, you can filter results by location and date after you have conducted a search. The most concerning area of this service is their Industry Control Systems (ICS) search page at https://www.zoomeye.org/topic?id=ics_project which allows the exploration of various online systems that control public utilities and various corporate infrastructure. Providing a target IP address into the search option at this resource will work very similarly to Shodan's search feature.

ThreatCrowd (threatcrowd.org)

ThreatCrowd is a system for finding and researching artefacts relating to cyber threats. Searching an IP address can reveal an association to malicious software being spread over the internet. A positive result will display the type of malware, associated domain names, dates of discovery, and any comments by other researchers. Most readers that actually need this type of service likely already know more about it than me. However, it should be a consideration when investigating suspicious IP addresses.

Censys (censys.io)

Censys is a search engine that enables researchers to ask questions about the hosts and networks that compose the internet. Censys collects data on hosts and websites through daily scans of the internet, in turn maintaining a database of how hosts and websites are configured. Researchers can interact with this data through a search interface. As an example, as search of 173.189.238.211 reveals it to be associated with a Schneider Electric BMX P34 2020 device through a Windstream provided internet connection, located near Kansas City, Kansas.

IntelTechniques IP Address Search Tool (inteltechniques.com/intel/osint/ip.search.html)

Similar to the domain tool mentioned previously, this page automates some of the most common IP address searches. The first box accepts any IP address. Clicking the Populate All button will insert this address into all of the search options where manual queries can be conducted. The final option will open several tabs within your browser that present each query listed on the page. Figure 17.03 displays the current status of the service.

Email Headers

I no longer teach email header analysis in my live courses. The vast majority of users rely on web-based email such as Gmail or Yahoo. These services do not disclose the IP address of an individual user within the email headers. The only email headers that I have encountered over the past three years that contained valuable IP addresses were business users that sent emails within a desktop client such as Outlook. If you would like to analyze an email header in order to identify the IP address and sender information, you have two options. You can look through a few sites and teach yourself how to read this confusing data, or you can use an automated service.

IP2Location (ip2location.com/free/email-tracer) provides a large text box into which an entire email header can be copied for analysis. The response includes the IP address and location of the sender, interactive map identifying the originating location, internet service provider, and links to additional information from an IP search. Anyone wanting more information from an email threat should start here. An alternative site that conducts similar actions is **MX Toolbox** (mxtoolbox.com/EmailHeaders.aspx).



Figure 17.03: The IntelTechniques Custom IP Address Search Tool.

Obtaining a Target's IP Address

You may want to know the IP address of the person you are researching as provided by their internet service provider. This address could be used to verify an approximate location of the person; to provide law enforcement details that would be needed for a court order; or to determine if multiple email addresses belong to the same subject. All of those scenarios will be explained here while I explain the various services that can be used.

What's Their IP (whatstheirip.com)

For many years, this was my favorite option for identifying the IP address of a target. There are many options now, and all of them will be explained here. This specific technique involves some trickery and the need to contact the target from a covert account. For this demonstration, assume that your target has a Facebook page that he checks regularly. You can send him a private message that includes "bait" in the form of an online link. Before you send the message, you must create the link. Navigate to whatstheirip.com and provide your email address. The target will not see this address, but I recommend an anonymous account. This will generate two website links that are unique to you. I always prefer the first option, as it looks less suspicious. Two links issued to me are as follows.

http://www.bvog.com/?post=IDAftMfYZQx9Sj7rp
http://www.hondachat.com/showthread.php?t=IDAftMfYZQx9Sj7rp

If a person visits the first link, they receive a notification that the website is no longer available. It is basically a blank page with an error at the top. The second link connects to what appears to be an online forum about cars, but the post requested is unavailable. Both of these links are designed to make the target believe that whatever was on these pages previously is no longer available. Your goal is to get your target to click on one of these links. When he or she does, this service will capture their IP Address and forward it to you. The easiest execution would be to email them a link. However, do not just send the link and hope for the best. I recently had a Craigslist investigation where a suspect was selling a stolen iPad through the website. I sent an email from a covert account that was verbatim as the following example.

Hi. I saw your ad on craigslist. I want to buy that iPad for my dad. I have cash and I live about 20 minutes away. I just need to know if it is the model 1 or 2. This link has a picture of what the back should look like. If it does, let me know and I can bring you cash today.

http://www.bvog.com/?post=IDAftMfYZQx9Sj7rp

When the target clicked on the attached link, he saw an error message stating "Not Found". Within one second of him clicking the link, I received the email visible in Figure 17.04. I immediately knew an approximate location and the IP address of the suspect. A subpoena to the internet service provider verified an actual address.

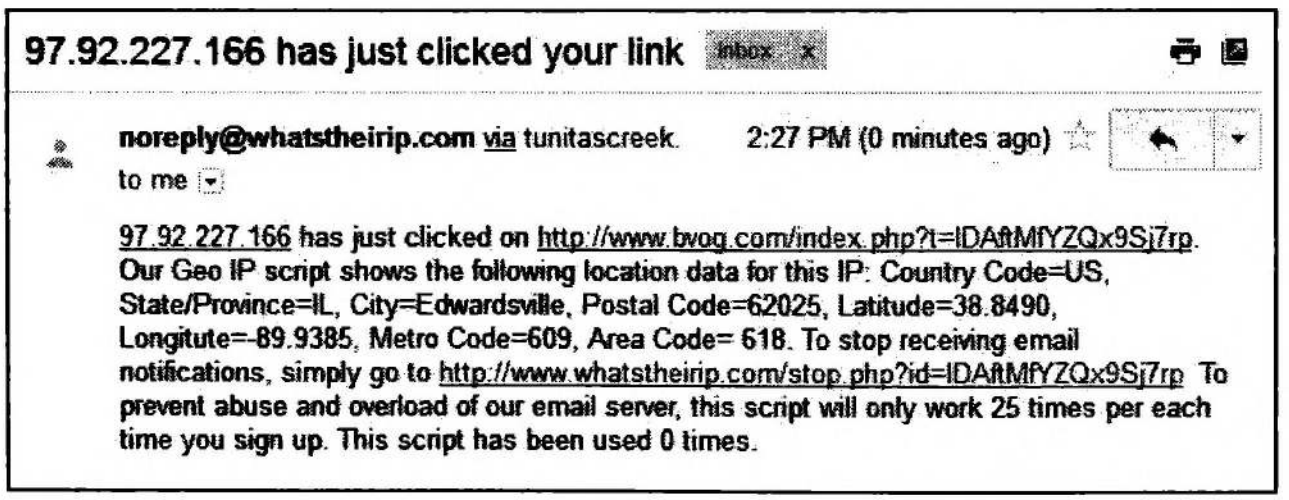


Figure 17.04: A response from "whatstheirip.com" identifying a target's IP address.

IP Logger (iplogger.org)

An alternative to whatstheirip.com is IP Logger. This service applies the same concept as whatstheirip.com but works differently. Instead of sending you an email when a target clicks a link, you must view a log file within the IP Logger website. Additionally, this service supports custom images that will track your target for extra covert privacy. A detailed set of instructions

should explain the processes with several options. The main website presents several options, but only the "Short Link" and "Your Image" services will be explained.

Link: You can generate a URL which will redirect to any website that you provide. IP Logger will save the IP address of each user who clicked the link. In the box provided, enter any address that you want the target to see when clicking on a link. This could be something generic such as cnn.com. After submitting, you will receive a series of links. This page also serves as the log of visitors, and I recommend documenting it. In an example, I received the following link at the beginning of this list.

http://www.iplogger.org/3ySz.jpg

Although the link appears to be a jpg image, clicking this link or typing it into a browser forwards the target to cnn.com. This action collects his or her IP address, operating system, and browser details. These details, along with the date and time of capture, can be viewed at the link generated previously. A URL shortening service such as Bitly (bit.ly) would make the link look less suspicious.

Image: You can provide a digital image to this service, and it will create a tracker out of it for placement onto a website, forum, or email message. I provided an image that is present on my website at inteltechniques.com/img/bh2016.png. This presented a page similar to the previous example. I was provided the following links.

http://www.iplogger.org/23fq.jpg <img src="http://www.iplogger.org/23fq.jpg">

The first link forwards to the image that I provided. During this process, the IP address, operating system, and browser details are collected and stored on the page that stored the links. The second link could be inserted directly into a web page or email address. Upon loading either, the image is present and collects the same data. I was once communicating with an unknown subject about illegal matters on a web forum about hacking and stolen credit card numbers. I wanted to find out his IP address in order to discover his true identity with a court order. I told the hacker that I had an image of a freshly stolen debit card that I was willing to share. He requested proof, so I created an IP Logger link based on a generic online image, and embedded that link into the web forum where we were communicating. Within a few moments, I visited the log for this image and discovered his IP address in Newark, New Jersey.

Blasze (blasze.tk)

A newer option for IP identification is Blasze. At the time of this writing, it was more popular than the previous two methods. It works very similar to What's Their IP. The difference is that you can forward the target to a desired link that may not raise any suspicion. As an example, assume that you want to send your target an email message that will identify his IP address. You want him to click on a link, but you do not want him to receive an error similar to what is

presented by whatstheirip.com. Instead, you want him to actually navigate to a safe website in order to eliminate any concern that he was compromised.

The Blasze website will ask you to enter a real website that you want your target to see. In this example, I will use the Reddit Netsec page located at reddit.com/r/netsec. When I supplied this link to Blasze, it generated a unique internet address (URL) of blasze.tk/DQ7ORY. This is the page that you want your target to open. It will forward him to the Reddit website, but it will first capture his information. Blasze will also generate a web page that will allow you to monitor the captured details. You should bookmark or save this link. In this example, the address for monitoring the forwarding link was http://blasze.tk/track/JYHFLR.

If your target clicks on the blasze.tk/DQ7ORY link, Blasze will track the information from his internet connection. However, this suspicious URL may make your target skeptical. Before the link is sent, I recommend using a URL shortening service that will make the link appear more trustworthy. I prefer to use Google for this. Navigate to the website goo.gl and enter the Blasze link provided to you (blasze.tk/DQ7ORY in this example). Google will generate a new link that will appear similar to http://goo.gl/dIviMz. Now, you can send a link that is less suspicious looking.

In this example, you would send your target the link of goo.gl/dIviMz. When clicking this link, it automatically forwards to the Blasze service and connects to blasze.tk/DQ7ORY. When this link is executed, it automatically forwards to the original "safe" site of reddit.com/r/netsec. Overall, the target clicks on a shortened Google link and sees the Reddit page. If this were my investigation, I would have sent a message similar to the following.

Hi David. Sorry, you don't know me, but I thought you should know that the project that you have been working on has leaked and is currently being discussed on Reddit here: goo.gl/dIviMz.

Obviously, you would want to use an anonymous email account. If your target opens the link, and sees the Reddit page, it will likely create confusion, but will look less suspicious than opening a link that generates an error message. You can then navigate to the Blasze link monitoring page, http://blasze.tk/track/JYHFLR in this example, and see the results. The data below displays my test results. The target clicked on the link at 22:30 hours on February 2, 2015. He was using a Chrome web browser (version 40.0), had an IP address of 68.225.11.142, and uses Cox as an internet service provider (ISP). IPLocation reports the residential IP address to be in Irvine, California.

2015-2-11 22:30:30 68.225.11.142 Chrome/40.0.2214.111 ip68-225-12-142.pv.oc.cox.net

You now know that there are several options for obtaining the IP address and general location of your target. New services such as **Grabify** (grabify.link) and **Canary Tokens** (canarytokens.org) offer redundant functionality as the previously mentioned products, but may be more useful to you. Ultimately, you should familiarize yourself with all of them and choose

which works best for you. Lately, I have found Canary Tokens to be the superior option of all. It allows creation of a PDF or DOCX file that contains a tracker, and is the most user-friendly of the services. After choosing a tracking option, it walks you through the process. I maintain a few Canary Token files at the following address. They are used as traps for people that conduct Google searches attempting to find my home address. Opening any of these alerts me to your IP address and general location. At the time of this writing, the most recent opening of one of these documents occurred only two days prior. The culprit lives in Matawan, New Jersey, possesses MCI as an internet provider, and had recently downloaded an Xbox 360 game through a torrent.

https://inteltechniques.com/canary

Always remember that technologies such as VPNs, Tor, and other forms of IP masking may create inaccurate results. You may choose to send these links from a spoofed email account. If I were to email you from mikethehacker@gmail.com and request that you open a link, you would likely delete the message. However, what if the email came from someone with whom you worked? It is easy to control the display of a sender's email address and name. There are many software applications that will allow this manipulation. However, the easiest way is through an online service.

Emkei (emkei.cz)

The service will allow you to immediately send an anonymous email message from within the website. You can completely control the sender's name, email address, subject, and message. The recipient will receive an email as normal. However, the "From" section will contain any spoofed information that you desire. Combining this utility with the tracking services mentioned previously can increase the success of the methods. Always test this technique by sending a message to yourself first. Secure email services such as Gmail will definitely flag these messages as suspicious. However, I have found many corporate networks that do not notify the receiver that the sending address appears fraudulent.

Anonymous Email (anonymousemail.me)

After noticing that anonymous messages from Emkei were being marked as spam by Gmail, I began using the free service from Anonymous Email to send messages that appear to be coming from someone else. I have found that Gmail does not always identify these as spam. Again, you should test any services in a controlled environment before actual execution.

Social Network

If you do not know your target's email address, you can send this same link to them through their social networks, such as Facebook. In September of 2013, I was investigating an incident that involved an anonymous Facebook profile that was harassing several people with violent threats.

A search warrant to Facebook would get what I needed. However, that can take several weeks. Instead, I sent a private message to the Facebook profile with the following text.

I don't know who you are, but I thought you should know that another Facebook user has posted your home address over on another site. You may want to take it down or at least be aware: http://www.bvog.com/?post=IDAftMfYZQx9Sj7rp

Obviously, this private information was not present. However, it was enticing enough to make the target click the link. That was all I needed in order to identify the IP address and approximate location of the target. If your person of interest is very tech savvy, he or she will know about this trick. If that might jeopardize your investigation, avoid this technique.

URL Biggy (urlbiggy.com)

If you want to add some flare to an IP logging link, consider this service. You can supply a link created with the previous instruction, but specify a new custom forwarding link. In other words, I can convert bvog.com/?post=IDAftMfYZQx9Sj7rp into urlbitty.com/michael-bazzell-home-address. This could entice the target to click a link that may otherwise appear suspicious.

BananaTag (bananatag.com)

This premium service offers a free version that is limited to five emails per day. It requires you to use a Gmail account and install a plugin through your web browser. In my tests, antivirus companies did not alert on the process. An occasional test on your end should be conducted if you continuously use this service. After you create your account and install the plugin as directed on the website, you will have a new feature within your Gmail account. Next to the standard "Send" button visible when you compose a new email, you will see a button titled "Track & Send". Clicking this button will inject a small piece of code into your message. When the message is read, you will receive an email announcing the event. You will need to login to your BananaTag account to see the details. Figure 17.05 displays details of the following email that I sent to my supervisor to determine his location (sorry boss).

LT, disregard that last message, I figured it out. MB

The response identified his IP Address, the operating system of his computer, the type of computer, the web browser, and the approximate location. This information tells me that he is likely at home and not replying from a mobile device. A message could be sent to multiple email addresses of your investigation. Changing each message may convince a single person using multiple accounts to click each link. This can verify that the multiple accounts actually belong to the same target.

My complaint with BananaTag is the number of steps involved to access the service. When conducting techniques like this, I prefer to use methods that do not require registration and the

installation of software. If this is something you will do on a regular basis, it is worth the configuration time. If you only need to identify one or two IP addresses, I recommend "Whats Their IP" and a bit of creativity.

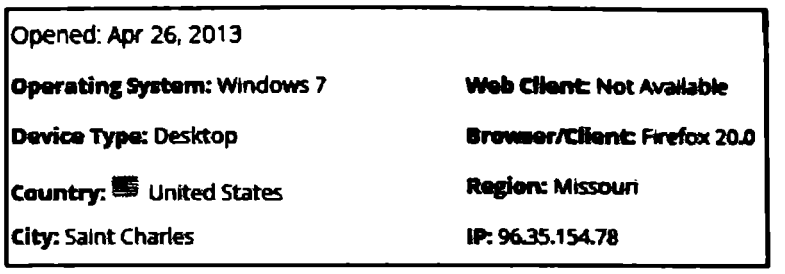


Figure 17.05: A BananaTag response.

Get Notify (getnotify.com)

Similar to BananaTag, Get Notify tracks the opening of email messages and presents the connection information of the target. However, this service is completely free and does not require Gmail as your email provider. You will need to create an account through the Get Notify website and you will be limited to five email messages per day. After you have registered the email address that you will be using, you can send emails from that account as usual. However, you will need to add ".getnotify.com" after each email recipient. Instead of sending an email address to the valid account of Michael@inteltechniques.com, you would send the message to Michael@inteltechniques.com.getnotify.com. This will force the email message to go through Get Notify's servers and route the message to the valid address. When your target reads the email message, Get Notify will track the user's IP address, geographical location, and notify you whether your message was viewed for a length of time or deleted right away.

Get Notify works by adding a small invisible tracking image in your outgoing emails. When your email recipient opens your message, this image gets downloaded from a Get Notify server. Get Notify will know exactly when your sent email was opened and it notifies you through an email that your sent message is read by the recipient. You can also view log files within your online account. The tracking image inserted by Get Notify is invisible to the recipient. Optionally, you can specify your own images to be used as tracking images by going to the preferences section after signing in to GetNotify.com. Your recipient will not see ".getnotify.com" at the end of his or her email address. If you want to send a single email to multiple recipients, you should add ".getnotify.com" at the end of every email address.

There are countless scenarios that may make these techniques beneficial to your online research. While I used it for law enforcement, especially in tracking down stolen goods on Craigslist, civilians can use it for many different things. Private investigators have used it on dating websites while hunting cheating spouses. Singles have used it to verify that the potential mate that they have been chatting with for weeks is really local and not in another state or country. The possibilities are endless.

CHAPTER EIGHTEEN
GOVERNMENT RECORDS

Open source government information has never been easier to obtain. A combination of a more transparent government, cheaper digital storage costs, and widespread broadband internet access has placed more information online than ever before. There is no standard method of searching this data. One county may handle the queries much differently than another county. The following resources and techniques should get you started in the United States.

County General Records (www.blackbookonline.info/USA-Counties.aspx)

Counties all over America have digitized the majority of their public records and allow unlimited access over the internet. Searching for your county's website will likely present many information options. This can become overwhelming and it can be easy to get lost within the pages of the site. My preference is to use Black Book Online's free county public records page. It allows you to drill-down from state to county. The resulting page isolates all available records for viewing. As an example, I chose Illinois and then Madison County as my target. I was presented with the following databases, each linking directly to the source.

Coroner Reports
Delinquent Tax Sale
Government Expenditures
Property Tax Search
Public Employee Salaries
Recorded Documents
Registered Lobbyists
Press Releases
Voter Registration Verification
Voter Registration Address Search
Unclaimed Property
Crime Map
Building Contractors
Building Permits
Foreclosed Properties

County Court Records (www.blackbookonline.info/USA-County-Court-Records.aspx)

A Google search of your county of interest should identify whether an online court records database is available. As an example, St. Clair County in Illinois possesses a website that has their entire civil and criminal court records online (circuitclerk.co.st-clair.il.us/courts/Pages/icj.aspx). Searching only a last name will present profiles with full name, date of birth, physical identifiers, case history, fines, pending appearances, and more. Navigating the website will expose charged crimes even if they were dismissed. This can be extremely useful in civil litigation. There are several websites that help connect you to publicly available county government records, such as Black Book Online. It allows you to drill-down to your local records. The main page will prompt for the state desired. The result will be a list of links that access each county's court information. Some rural areas are not online, but an occasional search should be done to see if they have been

added. Repeating my previous search of Madison County, Illinois revealed the following court related databases.

Circuit Court Complete Docket	Traffic Citations
Circuit Court Attorney Docket	Crash Reports
Family and Civil Pro Se Dockets	Police Blotter
Felony State's Attorney Jury Trials	Daily Crime Log
Traffic, Misdemeanor, DUI Docket	Jail Inmate Search

If the Black Book Online options do not provide optimal results, please consider **Public Records Online** (publicrecords.onlinesearches.com).

PACER (pacer.gov)

PACER is an acronym for Public Access to Court Electronic Records. It is an electronic public access service of United States federal court documents. It allows users to obtain case and docket information from the United States district courts, United States courts of appeals, and United States bankruptcy courts. As of 2013, it holds more than 500 million documents. PACER charges $0.10 per page. The cost to access a single document is capped at $3.00, the equivalent of 30 pages. The cap does not apply to name searches, reports that are not case-specific, and transcripts of federal court proceedings. Account creation is free and if your usage does not exceed $15 in a quarter, the fees are waived. I have possessed an account for several years and have never been billed for my minimal usage. PACER has been criticized for being hard to use and for demanding fees for records which are in the public domain. In reaction, non-profit projects have begun to make such documents available online for free.

RECAP (courtlistener.com/recap)

RECAP (PACER backwards) allows users to automatically search for free copies during a search in PACER, and to help build up a free alternative database at the Internet Archive. It is an extension for the Firefox and Chrome browsers which for each PACER document first checks if it has already been uploaded by another user to the Internet Archive. If no free version exists and the user purchases the document from PACER, it will automatically upload a copy to the Internet Archive's PACER database. While the browser extension assists greatly with searching, a search page exists on RECAP at the address above.

State Business Records (opencorporates.com)

Practically every state offers a searchable database of all businesses created or registered within the state. This will usually identify the owner(s), board members, and other associated subjects. **Dun & Bradstreet** (dnb.com) offers a business search within the home page, but many registered companies do not participate with them. In my experience, the best overall business lookup entity is Open Corporates. This free service indexes all 50 states in the U.S. plus dozens of additional

countries. The records usually identify corporate officers' names, addresses, and other contact details. The basic search option allows queries by business name, entity registration number, or officer name. Clicking the advanced option allows query by physical address, but requires you to create a free account. This website is superior to targeted Google queries, because it indexes and scrapes data directly from government websites. This can assist with identifying historical results that no longer appear within the original source. I visit this resource every time I encounter a business name during my research, or identify a target that would likely be associated with an organization.

Birthday Database (birthdatabase.com)

This site should identify the full name, date of birth, and city and state of birth of your U.S. target. The only available search fields are first name, last name, and approximate age. The age field is not required, but may help eliminate multiple results.

SSN Validator (ssnvalidator.com)

A simple way to verify if a number is valid is at SSN Validator. This does not provide the personal information attached to the number, only verification that the number is valid. A typical response will include the state that issued the number, the year issued, verification that the number was assigned, and confirmation of death if applicable.

Social Security Death Index (genealogybank.com/gbnk/ssdi)

This public index of death records is stored on a genealogy site. The only required information is the first and last name. The results will identify birth year, death year, state of last residence, and state of SSN issue.

Legacy (legacy.com)

There are many websites that search for death related information such as social security indexes and ancestry records. A leader in this area is Legacy. This site indexes online obituaries and memorials from approximately 80 percent of all online newspapers. The search on this site is straightforward and results can identify family members and locations.

Asset Locator (www.blackbookonline.info/assetsearch.aspx)

Black Book Online's Asset Locator is the most comprehensive list of sources for the search of real estate, judgments, bankruptcies, tax liens, and unclaimed funds. This page will allow you to select the type of asset you are researching and the state of the target. This will then create a new page with all of the options for that state. It will provide direct links to the sites for a search of the target. This often includes online databases of public employee salaries, vehicle registrations, property tax records, and dozens of other categories.

Vehicles

Many people assume that information related to vehicle registration and licensing is only available to law enforcement through internal networks. While a full driver's license search and complete license plate query is not publicly available, a surprising portion of related data is online for anyone to view. The following methods will display all publicly available details.

VIN Place (vin.place)

VIN Place is a website that provides free access to vehicle purchase data. All information on this website is public information, and the data comes from dealerships and auto insurance companies referencing new vehicle purchases. You can search by real name or VIN to find vehicle purchase information, vehicle specifics, and fuel economy information. I submitted four unique names and received two positive responses. When a record is found, it is quite revealing. One of my examples included the following information, which has been redacted here for privacy.

Address: REDACTED
City: COLORADO SPGS
State: CO
Zip: 80906-6544
Year: 2010

Make: VOLKSWAGEN
Model: JETTA
VIN: 3VWRL7AJ6AM13xxxx
Trim Level: TDI
Style: SEDAN 4-DR

VIN Search

After converting a real name into a vehicle identification number (VIN), you may want to verify the details through another service. The following options allow you to enter any VIN and retrieve the year, make, and model of the vehicle associated. The first option will often display estimated mileage based on service records.

Vin Coderz (vindecoderz.com)
CarFax (carfax.com/processQuickVin.cfx0)
Check That VIN (checkthatvin.com)
Search Quarry (searchquarry.com/vehicle_records)

NICB VIN Check (nicb.org/theft_and_fraud_awareness/vincheck)

While the previous two searches will identify details about vehicles and their owners, they will not display any information about theft or salvage records. The National Insurance Crime Bureau (NICB) allows search of any VIN and will display two unique pieces of information. The VINCheck Theft Record will identify vehicles that have been reported stolen while the VINCheck Total Loss Records identifies VINs that belong to salvaged vehicles.

Cycle VIN (cyclevin.com)

VINs from motorcycles may not be searchable on standard VIN engines due to the amount of characters in them. Cycle VIN will display a year and make, as well as any indication that the VIN exists in its proprietary database. If it does, $25 will obtain title and mileage information. I only use this as a free resource for verifying motorcycle VINs to the correct year and make.

Vehicle Registration

Several free services identify the year, make, and model of vehicle after supplying the license plate registration. As a test, I submitted a vehicle's registration number which was displayed on a television show playing at the airport while I wrote this section. The result correctly identified the vehicle as a 2010 Dodge Avenger. While only a small piece of information, it works in conjunction with other search techniques. When I have vehicle registration to search, I use the following resources, in the order from most available information to least. After exhausting all of these searches, you should be able to obtain the VIN, make, model, year, engine, and style of the vehicle. These options will not provide the name of the owner.

Reverse Genie (reversegenie.com/plate.php)
Auto Check (autocheck.com/vehiclehistory/autocheck/en/search-by-license-plate)
Vehicle History (vehiclehistory.com/licence-plate-search)
Records Finder (recordsfinder.com/plate)
CarFax (carfax.com/processQuickVin.cfx)
Search Quarry (searchquarry.com/vehicle_records)
Free Background Search (freebackgroundcheck.org)

Progressive (progressive.com)

While not an official vehicle search, the insurance provider Progressive offers an interesting piece of information. I first learned about this technique from S.L., a member of my online OSINT forum. When you view the home page at progressive.com, you are prompted to request a free insurance quote. If you provide the zip code and address of any target, you receive a summary of the year, make, and model of all vehicles registered at that address. You can supply any random data besides the physical address and receive the results. This was likely designed to make the quote process more efficient and accurate, but investigators should appreciate the free utility.

Marine Traffic and Boat Information

There is an abundance of details available about global marine traffic within ownership records and real-time monitoring. **Marine Traffic** (marinetraffic.com) provides an interactive map that displays the current location of all registered ships and boats. Clicking on any vessel provides the name, speed, collection time, and destination. **Boat Info World** (boatinfoworld.com) allows search of a boat name and provides the following details.

Boat Name
Boat Owner
Record Date
Registered Address
Hull ID
Hailing Port
Lloyd's Registry Number
Call Sign
Coast Guard Vessel ID
Service Type
Boat's Length
Boat's Gross Tons
Vessel Build Year
Ship Builder
Hull Shape
Propulsion Type

Aircraft Information

Monitoring aircraft during flight and searching historical ownership records is relatively easy. Commercial planes constantly announce their location with automated reporting systems and tail numbers act similarly to a vehicles registration plate. Today, this information is publicly available on multiple websites. **Plane Finder** (planefinder.net) displays an interactive global map identifying all known aircraft currently in flight. Hovering a selection displays the carrier, flight number, originating departure, destination, speed, and altitude. Historical ownership records are available on multiple websites and none are completely accurate. I recommend **Black Book Online's** aviation page (blackbookonline.info/Aviation-public-Records.aspx). At the time of this writing, it provided direct links to the following databases.

Aircraft N Number Search
Aircraft Ownership Search
Airline Certificates
Airport Profiles
Certified Pilots
Cockpit Voice Recorder Database
Flight Tracker
Military Aviation Crash Reports

Campaign Contributions

Any contributions to political campaigns are public record. Searching this is now easy thanks to three separate websites. These sites will search with information as minimal as a last name. Including the full name and year will provide many details about the target. This includes occupation, the recipient of the contribution, the amount, the type of contribution, and a link to the official filing that contains the information. After an initial search is conducted, you will receive additional search tabs that will allow you to filter by zip code, occupation, and year. Melissa Data allows you to search a zip code and identify all political donations for a specified year. The results from these sites may be redundant, but often contain unique data.

Open Secrets (opensecrets.org).
Money Line (politicalmoneyline.com)
Melissa Data (melissadata.com/lookups/fec.asp)

Criminal Information

If a target has a criminal past, there is probably evidence of this on the internet. County court searches will identify most of this information, but this requires a separate search on each county's

website. There are a handful of services that attempt to locate nationwide information by name.

Family Watch Dog (familywatchdog.us)

This is one of the leading sites in identifying public criminal information about sex offenders. The main page includes a "Find Offender" area on the left side. You can search here by address or name. The name search only requires a last name to display results. This will identify registered sex offenders that match the criteria specified. This will include a photograph of the target and details of the offense.

Felon Spy (felonspy.com)

This site can appear difficult to navigate at first. Most of the search fields for information forward to a sponsored result that will demand a fee for the information. The only free way to search this data is to click on the "Begin Search" button overlapping the map in the middle of the page. The only fields that should be searched on this page are in the top row and include address, city, and state. Entering any address in a target neighborhood will display markers on a map of convicted felons in that area.

Crime Reports (crimereports.com)

Crime Reports delivers a very comprehensive map of criminal incidents, traffic accidents, registered sex offenders, police reports, and emergency incidents. The only option for search is an address. Alternatively, you can move the map to a desired location. After you have selected an area of interest on the map, you can select the types of notifications to populate the map. These include violent crimes, property crimes, traffic issues, and emergency incidents. As long as you are only viewing a specific neighborhood and not an entire metropolitan area, you should be fine selecting all of the reports. This will mark all of the incidents on a map (Figure 18.01). These markers can be selected for further information about the incident. The marker will expand to display any details available about the incident. This often includes the date, time, address, crime, report number, and investigative agency. This type of detail can assist with an accurate filing of a Freedom Of Information Act (FOIA) request. Additionally, neighboring police departments can access data from another jurisdiction without a common report management system.

Inmate Searches

Both federal and state prisons offer prisoner details online. The amount of detail will vary by state, but most will include photographs of the target and details of the crime. In most states, this information is maintained in public view after the target is released, if the subject is still on probation or parole. Federal prisoners can be located at www.bop.gov/inmateloc. A first and last name is required for search. Each state maintains its own database of prisoner information. Conducting a search on Google of "Inmate locator" plus the state of interest should present official search options for that state.

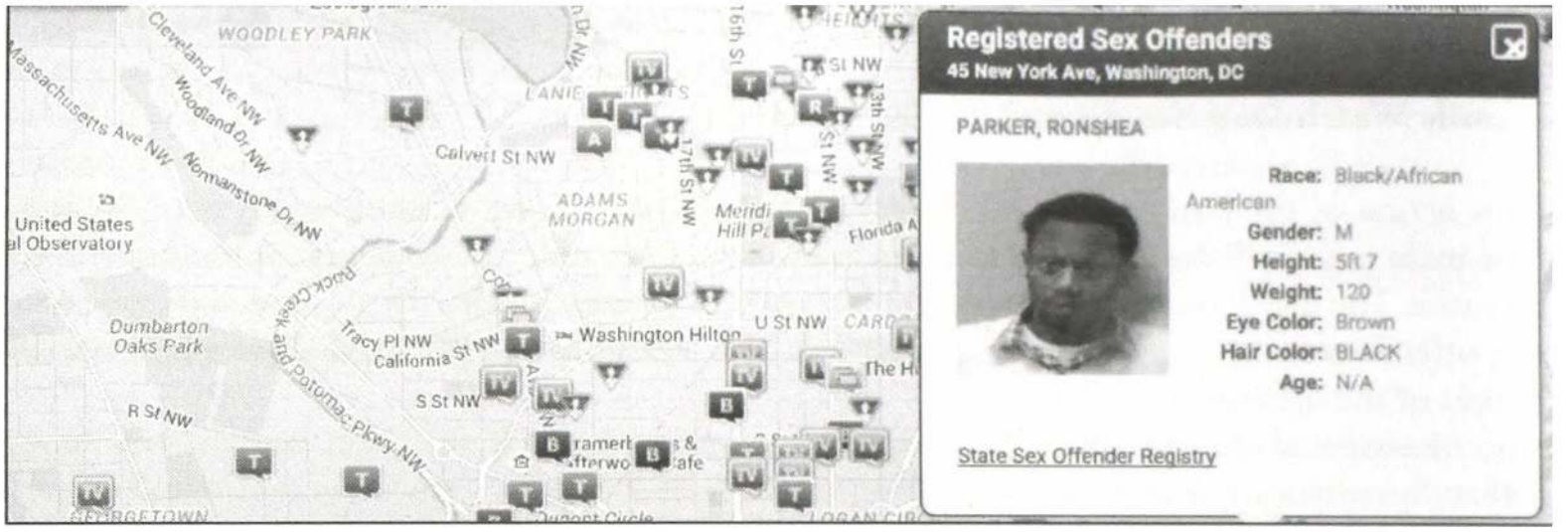


Figure 18.01: A Crime Reports detailed summary.

VINELink (vinelink.com)

VINELink is an online portal to VINE, a victim notification network. VINE has been providing victims and concerned citizens with information for decades, allowing individuals to access reliable information about custody status changes and criminal case information. After choosing the state of interest, you can select from the following options.

Find an Offender: Get info and register to be notified of custody status changes.
Find an Offender Court Case: Get info and register to be notified of offender court dates.
Find Sex Offender Registry Status: Get info about sex offender registry status changes.
Find a Protective Order: Get info and register to be notified of protective order status changes.

High Programmer (highprogrammer.com/cgi-bin/uniqueid)

Most states use some type of algorithm to create a driver's license number for a person. Often, this number is generated from the person's name, sex, and date of birth. After you have determined your target's middle initial and date of birth from the previous websites mentioned, you can use this data to identify the target's driver's license number. High Programmer will automate this process for the following states:

Florida	Michigan	New York
Illinois	Minnesota	Washington
Maryland	New Hampshire	Wisconsin

Military Duty Status (scra.dmdc.osd.mil/scra)

The website requires only a last name and date of birth in order to access a current active duty status report. This PDF document will open which identifies current status, leave of duty date, and future call-up date.

Selective Service Verification (sss.gov/Home/Verification)

This website requires a last name, social security number, and date of birth of the target. The result will identify the person's full name, selective service number, and date of registration.

Voter Registration Records

The election of 2016 caused a lot of controversy in regards to the use and collection of voter registration data. While these personal details are public record, many people did not believe it was appropriate for politicians to use this personal data as part of their campaign strategies. Regardless of your opinion on these matters, much of the voter registration details are available online. The most beneficial site I have found is at **Voter Records** (voterrecords.com). You can search by name or browse by state. Any results will identify full name, home address, mailing address, gender, party affiliation, age, and relatives. Currently, databases are available for Alaska, Arkansas, Colorado, Connecticut, Delaware, Florida, Michigan, Nevada, North Caroline, Ohio, Oklahoma, Rhode Island, Utah, and Washington.

Trace Checker (tracechecker.com)

Trace Checker is the largest database of property reported stolen to America's law enforcement agencies. It holds millions of serial numbers of stolen goods from thousands of police and associated agencies that can be searched for free by citizens and the police. It can help to avoid buying stolen goods or to identify goods that are recovered. You can check by serial number for any item of property. It will identify stolen goods as reported by the FBI Stolen Articles file on the NCIC database. You will first need to register for a free account. After you receive your login credentials, searching is easy. Enter the desired serial number in the "Property Search" area of the web portal. You will likely receive one of the following two results.

- Trace HAS NO records that indicate the serial number you have checked is associated with an item that has been reported as not being in the hands of the legal owner.
- Trace HAS records that indicate the serial number you have checked IS associated with an item that has been reported as not being in the hands of the legal owner.

BinDB (www.bindb.com/bin-database.html)

While not technically government data, I felt that this option fits best in this chapter. This website will allow you to enter the first six digits of any credit card number and identify the brand, issuing bank, card type, card level, country, bank website, and customer care line.

Real World Application: While working in the homicide division, I often identified credit or debit card numbers of my victims. If the actual card was located, I did not need this service. However, if only the number was located, this service helped to identify the financial institution and a contact number. In one specific investigation, I had learned that the victim had eaten at a

local restaurant the evening prior to her suspicious death. Visiting the restaurant allowed me to acquire the billing details of her dinner, which identified the debit card that she used for payment. Searching this number through BinDB identified the issuing bank and customer care number. Calling the number presented an automated self-service feature for members of that bank. Entering the newly found debit card number and the zip code of the victim allowed me to access the previous 30 days of charges to her account. This quickly identified an otherwise unknown ATM withdrawal on the day of her killing. Retrieving video from that ATM machine displayed a passenger in her vehicle. This initiated a new investigation which eventually led to the killer.

Bitcoin

I include information about Bitcoin in the Government Records chapter because it fit better here than anywhere else. In simplest terms, Bitcoin is digital currency, and can be spent for goods and services without connection to a person or bank account. It has no physical presence, and is mostly used online as digital payment. A bitcoin address, which is an identifier which you use to send bitcoins to another person, appears similar to a long string of random characters. In our demo, we will use 12t9YDPgwueZ9NyMgw519p7AA8isjr6SMw, which is the real address that was used to collect ransom from victims after malicious software had taken over their computers. Think of a Bitcoin address as an email address. That address stores their money.

Blockchain (blockchain.info)

This website allows search of a Bitcoin address and displays the number of transactions, total amount of Bitcoin received ($), final balance, and a complete transaction history. We can track every incoming and outgoing payment. This will almost never be associated with any real names, but provides a great level of detail about the account. We learn that this account has received 19.12688736 Bitcoin worth $ 287,391.14 USD at the time of this writing.

Bitcoin Who's Who (bitcoinwhoswho.com)

Our next stop is a service that provides a bit more analysis about the suspect account. We immediately learn that it is a suspect ransomware account, and that the address has appeared on various news outlet websites. Furthermore, we see transaction IP addresses, which are likely behind VPNs. Overall, I use Blockchain for transaction details and Bitcoin Who's Who to get a better idea of why I might care about the account.

Further Research

There are several websites that provide links to the numerous online government records. With over 3,000 counties in America, multiple sources should always be used. Some of the better collections can be located at the following websites.

brbpub.com/free-public-records publicrecords.searchsystems.net

CHAPTER NINETEEN
SOFTWARE APPLICATIONS

Up to this point, every technique that has been discussed will work on any computer regardless of operating system (OS). This includes Windows, Mac, and Linux. As long as the system has an internet connection and a web browser, you can perform all the methods for searching. This chapter focuses on software applications that can assist with your OSINT searching. Most of these programs will only work on a computer with a Windows operating system. This will include the majority of readers. If you have a Mac, you can use virtualization software to run Windows within it as explained in Chapter One. When available, I will explain Mac options within each program. As a reminder, the Buscador Linux Virtual Machine explained in Chapter Two possesses many of these options, pre-configured for immediate use.

Many of the programs detailed here can be executed as a "portable application". This means that the program can be downloaded and started without installation. All of the files needed are within the application and ready to go. This can be convenient when an application is needed on another computer. You can download the programs to a portable flash drive and execute them directly from the drive. Storing these applications on a portable drive will keep you prepared to conduct queries from any location. I recommend visiting **Portable Apps** (portableapps.com), **Pen Drive Apps** (pendriveapps.com), and **Portable Freeware** (portablefreeware.com).

All of the software mentioned in this chapter is free for personal use. Before using, read any help files included to ensure that you are not violating any of the terms of service. I have yet to come across any concerns. Most of these programs are updated regularly. These updates are very important. They often correct software bugs and add new features. Some, but not all, of these applications will notify you when an update is available. Visit the application website occasionally to see any updates.

All programs mentioned in this book are legal to use. They are tools that should be used responsibly. Some of them may be considered hacker tools, often used for malicious or devious purposes by an individual. It is my intent that the programs will be used only to obtain information legally. It is the responsibility of the reader to do so. Please be sure to only install these programs onto computers that you have the authority to do so. If you are a government or corporate employee, you likely have a policy that discusses computer usage at your workplace. This probably forbids you to place unauthorized programs anywhere within the network. Contact your I.T. representative before blindly downloading and installing any of these applications. To begin this chapter, I start with the software mentioned within Chapter Two, but will explain the usage for Windows and Mac instead of the Linux configurations.

Video Manipulation Utilities (ffmpeg.org)

When I was in law enforcement, one of my assignments was to obtain any video evidence from various crime scenes. This might be video footage of an armed robbery from a convenience store camera system or personal video captured through a cellular telephone camera by a witness. Either way, I often encountered many problems. Some surveillance systems required rare video codecs in order to view the media. Some personal videos were very short and difficult to see because of quick movement. I developed scripts to take advantage of open source tools that helped with the various struggles that I had involving digital video. The following techniques will require the executable files ffmpeg.exe and ffplay.exe. Downloading the correct version of FFmpeg can be difficult. Navigating to the above website will present dozens of options for various operating systems. In order to make this easier, you can type in the exact address below and download a working version for Windows. It also includes every script in this section.

https://inteltechniques.com/data/ffmpeg.zip

Video Codec Player

When you identify an online or offline video that offers value to your investigation, you should save and archive the file. This video, if gathered from a digital surveillance system, may require a video codec that you do not possess. If you cannot view the video, it has no value. I recommend using ffplay.exe as your best chance of viewing videos that require unknown codecs. You can play a video file by executing instructions via a command line. However, I prefer to create a batch file that will simplify the process every time. Extract the two files (ffplay.exe and ffmpeg.exe) from the compressed file mentioned previously. Save them inside a folder titled "video" on your desktop. In this same folder, create a new text file and title it player.bat. Be sure to change the file extension from txt to bat. Windows will now recognize this text file as a set of instructions. Type the following text into this new file.

```
set /p VIDEO=Video file name (with extension) on Desktop:
ffplay.exe "%userprofile%\desktop\%VIDEO%"
```

Double-click this new batch file and you should be prompted to enter the name of an unplayable video file. Note that you must supply the entire file name, including the file extension, and the video file must be placed directly on the desktop. You should now be able to play the previously unplayable file. Note that some video files will still not play. However, this method should eliminate many of the problems.

Video Converter

If you are able to now play the previously unplayable video file, you should consider converting it to a more universal format. During my investigations, I was often asked to forward any video evidence to a prosecutor. If I had a hard time playing the video, it was certain that the prosecutor

would have difficulty. Therefore, I always submitted both the original evidence video and a converted copy that should play on any computer. We can use the ffmpeg.exe file to convert any playable video to a standard MP4 format. Create another text file within your video folder and title it converter.bat. Type the following text into the document and save it.

```
set /p VIDEO=Video file name (with extension) on Desktop:
ffmpeg.exe -i "%userprofile%\desktop\%VIDEO%" -vcodec mpeg4
"%userprofile%\desktop\%VIDEO%.mp4"
```

Double-click this new batch file and you should be prompted to enter the name of an unplayable video file. Note that you must supply the entire file name, including the file extension, and the video file must be placed directly on the desktop. This should create a new video file on your desktop that will have the same title as the previous file but with the extension MP4. This file should play on any modern computer system.

Video Frame Extraction

You may also want to extract the still frames from the video for deeper analysis. Law enforcement may want to extract the stills of online videos for distribution to the media with the hope of identifying a suspect. There are many expensive programs designed to offer a solution for this, but this free program works just as well. The ffmpeg.exe file downloaded earlier will extract still images from practically any video and save them as uncompressed bitmap (BMP) files. Create another text file within your video folder and title it extract.bat. Type the following text into the document and save it.

```
set /p VIDEO=Enter full name of video file on desktop:
md "%userprofile%\desktop\frames"
ffmpeg.exe -y -i "%userprofile%\desktop\%VIDEO%" -an -r 10
"%userprofile%\desktop\frames\img%%3d.bmp"
```

It is very important that these files do not have a "txt" extension. If you double click this file and it opens within Notepad, the file extension is "txt" and not "bat". Change the file extension and you should see a black box open. In Figure 19.01, I had a video titled 1.dav on the desktop of my computer. After entering this file name of the video, press enter on your keyboard and you should see the program begin processing the video. When complete, you should see a new folder on your desktop titled "frames" which will contain numerous still images in chronological order. These are the frames from your video which can now be printed, distributed, or enhanced. If you have multiple videos to process, be sure to rename or remove the Frames folder before each new execution. Figure 19.02 displays a portion of the still frames available in the previous example.

Video Audio Extraction

During a law enforcement training event in 2015, a Detective asked if it were possible to extract

only the audio feed of a video file. He had several video files containing suspect interviews, and did not want to be stuck next to a computer in order to review the recordings. He believed there would be value in the ability to possess an audio file that could be played universally on any mobile platform. During the class, I created a batch file that used FFmpeg to extract the audio from any video file. Enter the following into a text file and save it as audio.bat within the same video folder created earlier. This script will extract the audio track from the supplied video and save it as a 320k MP3 file on the Desktop. The file name will be identical to the video file name with MP3 added to the end.

```
set /p VIDEO=Video file name (with extension) on Desktop:
ffmpeg.exe -i "%userprofile%\desktop\%VIDEO%" -vn -ac 2 -ar 44100 -ab 320k -f mp3
"%userprofile%\desktop\%VIDEO%.mp3"
```

As a reminder, all of these video manipulation scripts, as well as a recent binary of FFmpeg is available at https://inteltechniques.com/data/ffmpeg.zip. Unzip this file onto your computer to immediately execute any of these actions.

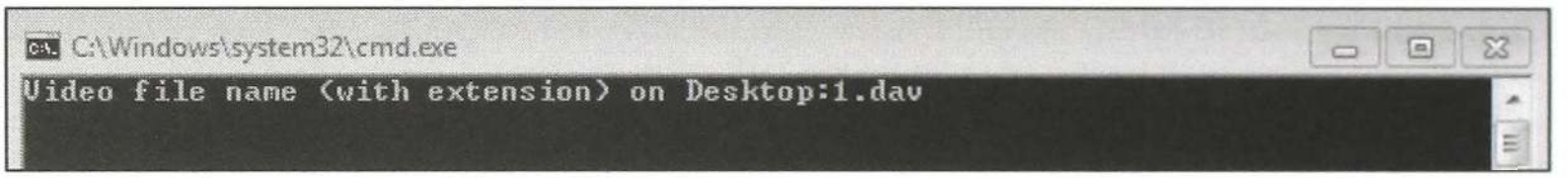


Figure 19.01: A batch file for video frame extraction.

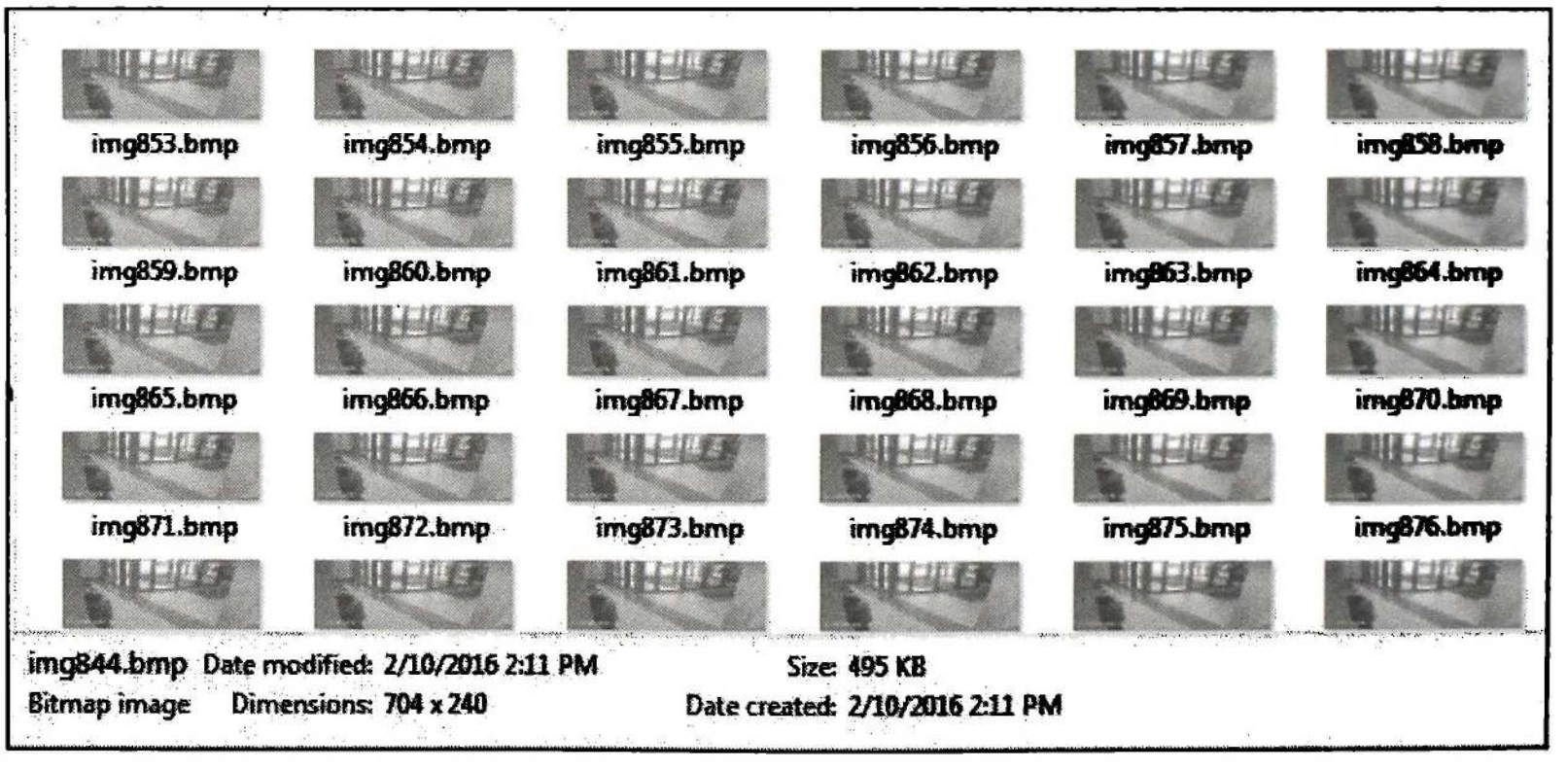


Figure 19.02: Still frame captures from FFmpeg.

Video Download (rg3.github.io/youtube-dl/download.html)

I mentioned YouTube-DL previously as my absolute favorite video download utility. In Chapter Two, I demonstrated how this Python script had been configured as a point-and-click option within the Buscador system that downloaded individual or bulk online videos from hundreds of websites. I rely on this tool so often that I have replicated its features for easy access on Windows and Mac computers. If you are using Windows, the above website offers a "Windows exe" download option. Mac users can obtain the latest release in bold to the left of "(sig)" on the website.

Windows

Save the downloaded executable (exe) file directly to your Documents folder. Double-clicking this file will not produce any desirable result. Instead, open the Command Prompt utility. This will be available in different locations within various versions of windows, but typing "cmd" in the run prompt should load the black terminal window. Once you have the Command Prompt application open, you need to navigate to the Documents folder. Execute the following command.

```
cd Documents
```

You can now launch a command that will execute the YouTube-DL application and any desired parameters. This was explained in Chapter Two, but the instructions are a bit different on Windows. Assume that you have located the same Bob Ross video channel that was discussed earlier (youtube.com/user/BobRossInc/videos). The following command would launch YouTube-DL and download all of the videos available in this channel in the highest quality possible. Below this command is the description of the actions taken.

```
youtube-dl.exe -f "best[ext!=webm]" --all-subs -o "%userprofile%\Desktop\Videos\
%%(title)s.%%(ext)s" --write-info-json  -i https://youtube.com/user/BobRossInc/videos
```

youtube-dl.exe: The command to execute the script
-f: Forces download of best available quality and ignores webm versions
--all-subs: Downloads text file of subtitles
-o: Specifies location of downloaded media
%%(title)s.%%(ext)s : Titles the video file with the video name and extension
--write-info-json: Download associated metadata
-i: Ignores any errors
https: The link of the target videos

While replicating this process every time that you need to download videos is not difficult, it can be time consuming. Therefore, consider creating an automated "batch" file that will provide point-and-click access to this utility. Open Notepad within Windows and type the following exact

text into it, saving the result as videos.txt in your Documents folder. Through File Manager, change the name of this file to videos.bat. You may need to disable "Hide extensions for known file types" in the Folder Options Window.

```
set /p VIDEO=Entire URL (Address) of video or channel page:
youtube-dl.exe -f "best[ext!=webm]" --all-subs -o .
"%userprofile%\Desktop\Videos\%%(title)s.%%(ext)s" --rm-cache-dir --write-info-json -i
%VIDEO%
pause
```

You can now double-click the videos.bat file and be presented with a Command Prompt ready to simply accept a URL of your target video page. Pasting the video address into this window executes the exact command required to replicate the process previously completed manually.

Mac

Similar to the Windows directions, save a copy of the proper YouTube-DL file in your Documents folder. Open TextEdit and type in the following exact text. Save the file as videos.command in the Documents folder.

```
#!/bin/bash
echo "Enter the Video or Channel URL: "
read youtube_url
python ~/Documents/Apps/Portable/YT-DL/youtube-dl.py $youtube_url -f best[ext!=webm]
-o ~/Downloads/"%(title)s.%(ext)s" -i
```

Open the Terminal application and execute the following commands

```
cd Documents
chmod +x videos.command
```

You should now have an executable file called videos.command in your Documents folder. Double-click on this and you will be prompted to enter the URL of your target video page. Any videos downloaded with this script will automatically be placed in your Documents folder. In order prevent typographical errors, I have created the two text files referenced previously and placed them on my website for easy access at the following links.

https://inteltechniques.com/data/windows.txt
https://inteltechniques.com/data/mac.txt

Video Metadata (mediaarea.net/en/MediaInfo/Download/Windows)

Most smartphone devices store data within every video captured. This includes the software version and model of the phone, the date and time of the video, and the GPS location of the device during the capture. It also documents the direction that the phone was facing during the capture as determined by the internal accelerometer. There are several ways of extracting this type of information. I have found MediaInfo to be the easiest solution. Navigating to this download website will present you with many options. Most users download the universal installer which will quickly install the required software. Unfortunately, it will also display advertisements and attempt to trick you into installing unnecessary software that is difficult to remove. My preference is to download the 32 bit "CLI" option. This is a command line version which will require a bit more work to make it run. However, you do not receive any unwanted bundled software.

Download the compressed file and unzip it to a folder called metadata. Create a new text file within this same folder and title it metadata.bat. Be sure that this file no longer possesses a txt file extension. Type the following text into this new batch file.

```
set /p VIDEO=Enter full name of video file on desktop:
mediainfo.exe "%userprofile%\desktop\%VIDEO%" > "%userprofile%\desktop\%VIDEO%".txt
```

Double-click the metadata.bat file and you should be prompted to enter the name of a video file residing on your desktop. Place any video file of interest on your desktop and enter that file name. This will create a new text file on your desktop that will have the same name as your target video. The content of this new "report" will contain all of the available information about that video from the metadata. The following partial information was retrieved from a test video that I extracted from a co-worker's cellular telephone.

Complete name:	C:\Users\Office\desktop\2.mov
Format:	MPEG-4
Format profile:	QuickTime
File size:	1.00 MiB
Duration:	10s 712ms
Overall bit rate:	787 Kbps
Recorded date:	2013-08-26T07:46:36-0500
©xyz:	+38.8890-090.1599+161.000/
Model:	iPhone 4S
Writing application:	6.1.3

This information identifies the location that target video was captured, the make and model of the device, and even the operating system within the iPhone. While many people realize that their phones record GPS within the photos that they take, they do not always know that this applies to videos as well.

Google Earth (earth.google.com)

Google maps is an online website that was discussed in Chapter Twelve. Google Earth is a standalone application that takes the Google Map data to another level. With this application, we have access to many mapping tools. These tools can import data from spreadsheets and help you visualize the content. In order to maintain the scope of open source intelligence, I will focus on only a few specific tools. Within the application, the first step is to display your location of interest. This can be accomplished by typing the address or GPS coordinates in the upper left search field. When you see your target location and have set the zoom to an appropriate level, you are ready to start adding layers. By default, you will only see the satellite imagery of the location. The menu on the left possesses options for adding new content to this view. The last box in this menu is titled "Layers". Inside of this menu are several data sets that can be enabled and disabled by the checkbox next to each. The following details will explain the layers of interest.

Photos - Digital images uploaded through social networking sites Panaramio and 360cities
Roads - Text layer of road names
3D Building - Alternative 3D view of some locations
Gallery - User submitted content including YouTube videos

I recommend disabling all layers and then enabling one at a time to analyze the data that is added to your map view. Figure 19.03 displays a view of Chicago including the Photos, Roads, and Gallery layers.



Figure 19.03: A Google Earth view with layers enabled.

Another Google Earth tool available that is often overlooked is the Historical Imagery option. This can be activated by selecting the "clock" icon in the upper menu bar of the application. This will open a slider menu directly below the icon. This slider can be moved and the result will be various satellite images of the target location taken at different times. Figure 19.04 displays the same target area with the Historical Imagery option enabled. The view has been changed to the

satellite image obtained on 05/30/2008. Usually, the quality of the images will decline as you navigate further back in time (Figure 19.05). This can be useful in identifying changes in the target location such as building modifications, additional vehicles, and land changes. Drug enforcement agents often use this tool to monitor suspected drug growth at a target location.

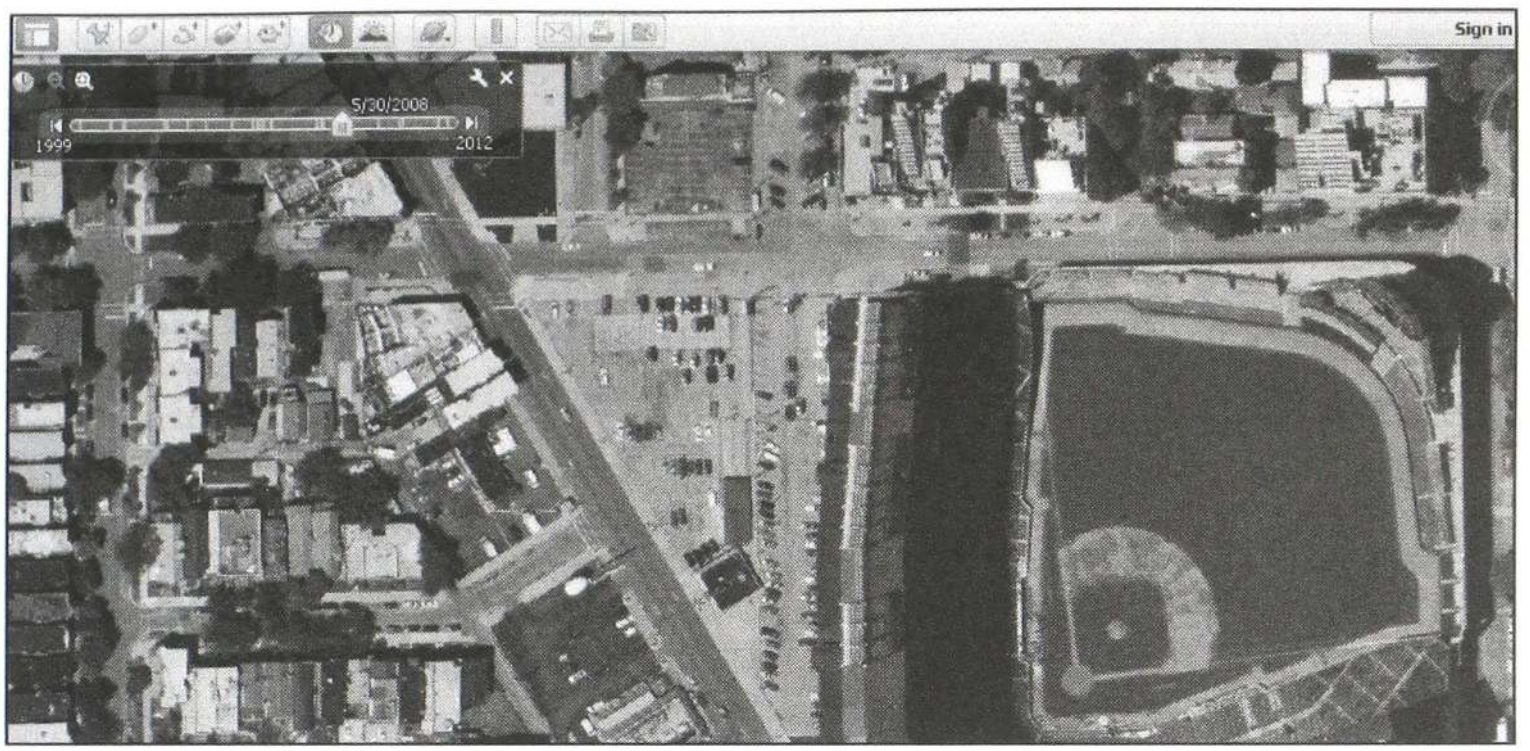


Figure 19.04: A Google Earth view of historic imagery from 2008.

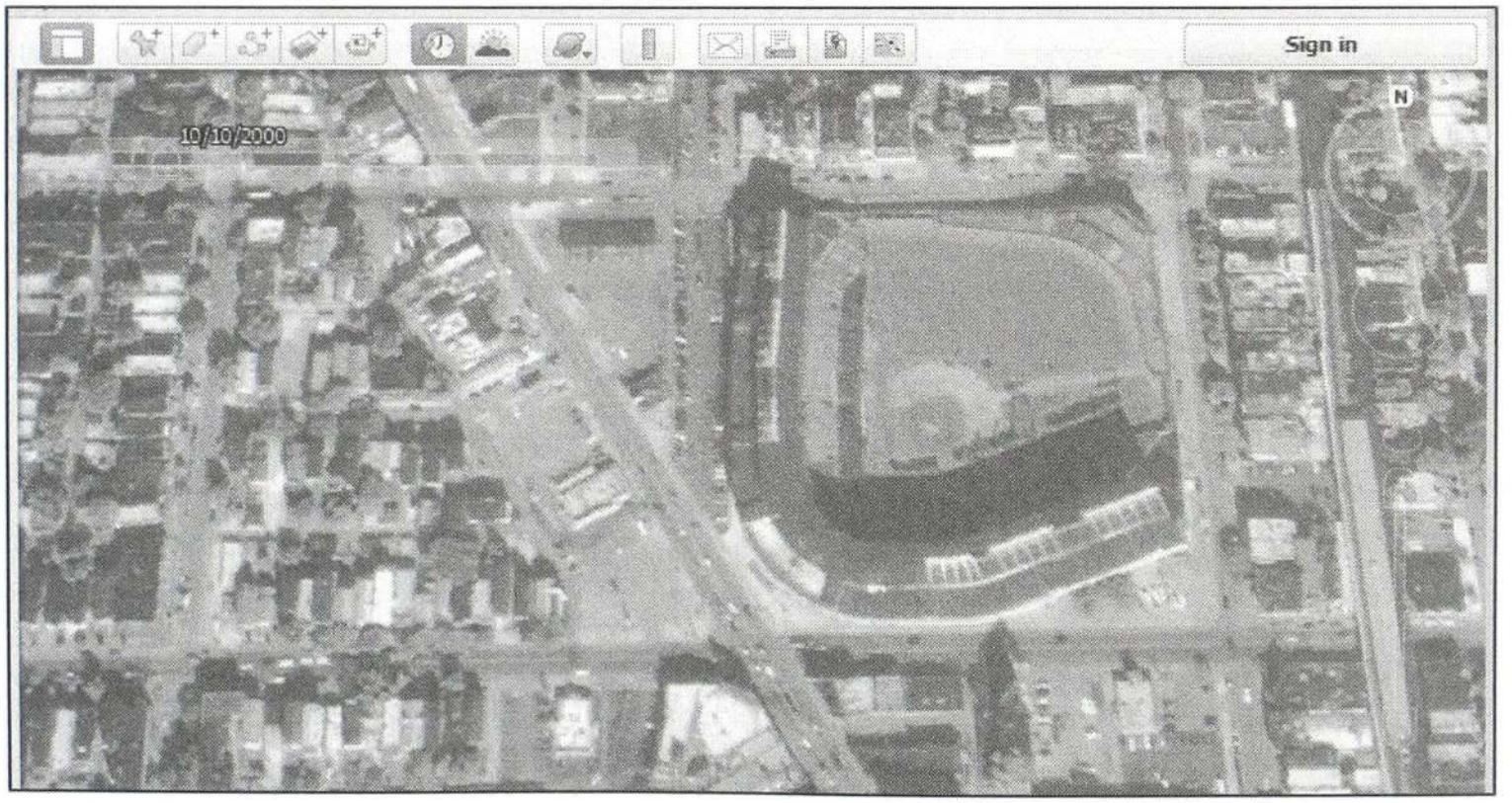


Figure 19.05: A Google Earth view of historic imagery from 2000.

Creepy (ilektrojohn.github.com/creepy)

This utility has been around since the second edition of this book. In its early years, it was an extremely valuable tool that I launched during almost every Twitter investigation. Today, its power has been crippled by Twitter's strict enforcement of their API limits. However, there is still value here for specific investigations. Creepy is an application created to display a Twitter user's location on a map, or Twitter users that have posted from a specific location. This is determined by the GPS data stored within a Twitter post. This can identify places visited by a target with the date and time that they were present. The program previously allowed searching of any combination of Twitter user names, Instagram user names, Flickr user IDs, or locations. However, it currently only works with Twitter data, as Instagram and Flickr have blocked the application by changing their own API rules.

Before attempting a search, click on "Edit" in the menu and then "Plugins Configuration". The Twitter plugin will ask you to log into your account and will guide you through the API setup process. Look for the button titled "Run Configuration Wizard". Create a new project and enter the user name of your target. Select the networks that you want to search and click "Search" to find the accounts. Any accounts identified will be displayed below. Click "Add to Targets" to select the accounts desired. Continue this process until you have added any accounts of your target. Accept the default options and click "Next" and "Finish" to start the query.

The application will identify posts that contain GPS information from the selected accounts. It will then map out each post on an embedded Google map. The column on the right will display all of the geo-located posts in chronological order (Figure 19.06). You can double-click any of them to see additional information. The map will change so that the center marker is the location of the chosen message. The lower right window will display the message and a link to the original source. The latest version will allow you to enter multiple targets from numerous accounts. Each project will automatically be saved within the application and available to you with the next launch. You can right-click any project to delete it. The program allows you to export a project to a standard CSV file or a Google Maps KML file. The KML option allows you to open the analysis within Google Maps.

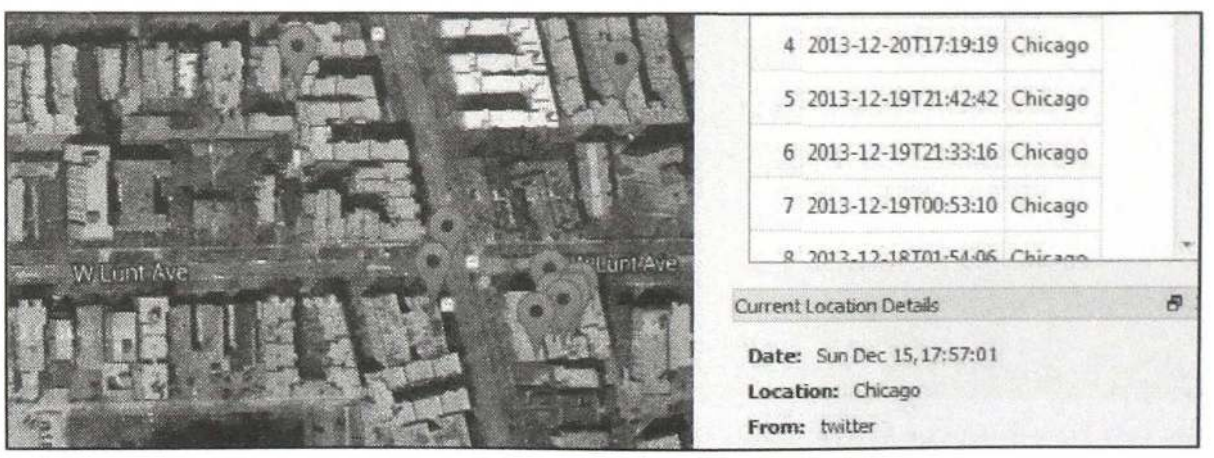


Figure 19.06: A Creepy Twitter result.

Exif Tool (sno.phy.queensu.ca/~phil/exiftool)

The details about how Exif data is stored within images was discussed in Chapter Fourteen. This data can be very valuable to an investigation. The resources mentioned earlier require internet access and the image to be uploaded to a website. For most situations, this is not a problem. Some investigations may involve classified material that is prohibited from being uploaded to any public network. Under these circumstances, Exif Tool becomes useful. The program is portable and allows you to browse to an image. It will then display all stored Exif information about the image. This will include camera details and GPS coordinates if available. Since this application does not use the internet for information, you will not receive a map of the location. You will only get the numerical coordinates.

JPEG Snoop (impulseadventure.com/photo)

If Exif Tool does not provide enough data, or displays a view of the data that you do not like, take a look at JPEG Snoop. This portable application provides a very detailed report of all of the Exif information that is stored. The report can be saved to a log file or printed. The best feature of this software is the batch process ability. You can select an entire folder of images and conduct an analysis of all of them at once. A report can be generated of the results, which can be archived to disk. JPEG Snoop will not display a map and does not require internet access (Figure 19.07).

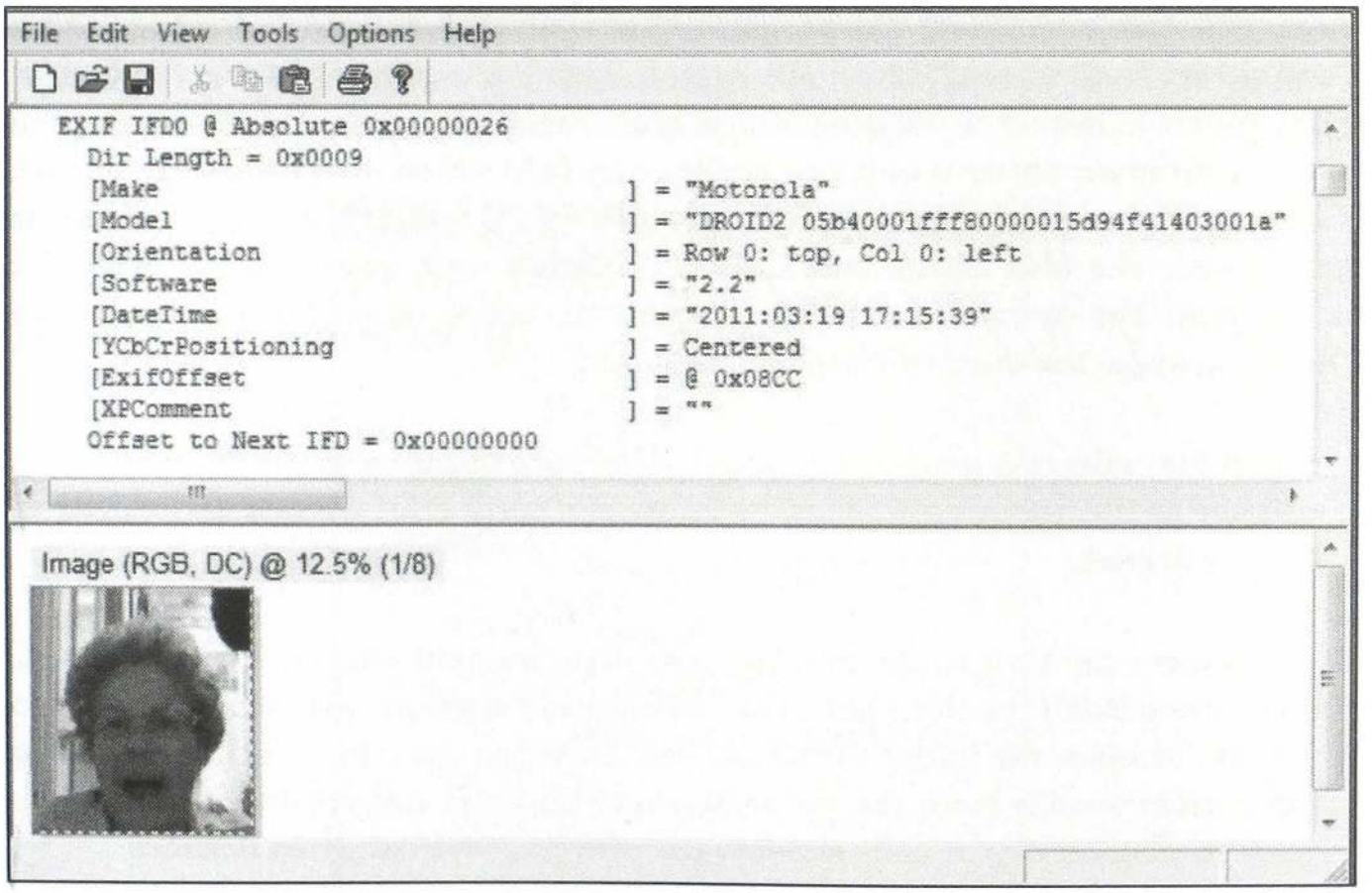


Figure 19.07: An image analysis in JPEG Snoop.

HTTrack (httrack.com)

There are several ways to make an exact copy of a website. I choose HTTrack because it is portable software and works quickly. The software will walk you through the process of saving a website. You may want to do this when you locate a target website and are concerned that the site could be taken down. Any time that you find any content that will be used in court, you should archive the entire site. This application automates the process. The result is a locally stored copy that you can navigate as if it were live. This is beneficial in court when internet access is not appropriate or a website has been taken offline.

When the application loads, clicking the "next" button will bring you to the project screen. Give your project a title, preferably the website name, and choose a location to save all of the data. Click next and then "add URL". Enter the exact website that you want to archive. Do not enter any login or password. Click "next", then "finished", and you are done. The application will begin extracting all public information from the site. This can take a while, depending on the size of the site. When complete, you will have an exact copy in which you can navigate and search offline. Archiving to CD or DVD creates a replica that can be held for future analysis.

4K Stogram (4kdownload.com/products/product-stogram)

4K Stogram, which I refer to as Instagram Downloader, is a standalone software application available for PC, Mac, and Linux. The program allows you to download and backup all Instagram photos and videos from a user at once. The default download option is a setup file that will install to your computer. However, a portable option is also available within the "Download" options. Launching the program presents only one single entry field which is ready for any Instagram user name. Clicking "Subscribe" will begin the extraction of all public photos and videos from the specified account. The files will be saved to your default documents directory in a subdirectory titled 4k Stogram. The options within this program are not obvious. The following instructions should help you make the most of the program.

- Allow the collection process to complete before navigating to the content folder. A grey counter in the upper right should identify the number of photos and videos that have been collected.

- Hover over the user name directly above the collected photos. This will activate three dots immediately to the right. This will present a menu that will stop the collection process or open the folder where the files are being collected. Click "remove" to delete the current search from the program view. This will not remove any files. You must remove all content manually through the operating system when desired.

- If you are seeking files from private accounts, and you possess an account that is "friends" with the target, you can enter your credentials within 4K Stogram to access the restricted account.

CamStudio (camstudio.org)

If all else fails and you cannot retrieve a pure copy of a video, you can always create a video screen capture while the video plays. This is not the recommended plan, but is better than not archiving the footage. CamStudio is completely free and simple to use. It can also work as a portable application from a flash drive. Upon launch, the program is ready to start recording. Before hitting the record button, I suggest visiting the "Video Options" menu under "Options". If you are going to use this recording as official documentation of an investigation, you should increase the quality to 100%. This will create a large file, but the quality is worth the size. Now, you must choose what "Region" you want to record. Your options are "Full Screen" or "Fixed Region". If you are working on a single monitor, "Full Screen" should work fine. If you have multiple monitors and do not want everything on them captured, you should select a region. This should be the only mandatory configuration changes. You are now ready to record.

Clicking the red record button will start the recording and place the menu icon in the taskbar. When you are finished recording your screen, right-click on the menu in the taskbar and choose "Stop". This will prompt you to title your video and choose a storage location. This video is now ready for archiving. This screen recording technique could also be used to capture an entire OSINT search and analysis. You could start the video before any searching is conducted and let it record while you navigate through websites. This could then be used for reference later to review the steps taken to locate any vital data. Some investigators like to archive this for court to confirm that the data obtained was indeed located legally and through open source methods. If you do choose to record your research, I encourage you to disable the "Record Audio" feature, which will prevent your microphone from recording your voice.



Figure 19.08: A CamStudio Options menu.

Lightshot Capture (app.prntscr.com/en)

If still captures are needed of anything on your screen, you have several options to generate them. You could capture a screen shot with the print screen button and paste it into a program that will allow you to save it. However, this is a hassle and time consuming. You could print the page you are viewing, but this gets difficult on some sites, such as Facebook. You could save the entire web page to your hard drive and recreate it as needed, but this is overkill. I recommend an automated screen capture. Lightshot Capture will do this for free. Some configuration is necessary, but since the application is portable, you can take your changes with you. Lightshot is available for Mac and Windows.

Launching the program will present a small icon in the taskbar. It will look similar to a purple feather. Right-click on this and choose "options". This will present several settings that you can customize for your needs. You can choose which combination of keys will create a new capture and what type of file is created. I use the PNG option and escalate the quality to 100%. You can now click on the Lightshot icon and your screen will go dark, waiting for you to select the portion that you want saved. Alternatively, you can hold the Shift button and the PrntScrn button to capture the entire screen automatically. If you select a portion of the screen with the default setting, you will be presented with a menu of options, as seen in Figure 19.09.

In order from lower-left to upper-right, the icons each perform the following tasks.

Upload screen capture to Lightshot's servers (NOT recommended)
Share screen capture on social networks (NOT recommended)
Conduct a reverse-image search on Google
Print the capture
Copy the capture to your clipboard
Save the capture
Delete the capture
Annotate the capture with color, text, highlighting, boxes, arrows, lines, or free-form writing

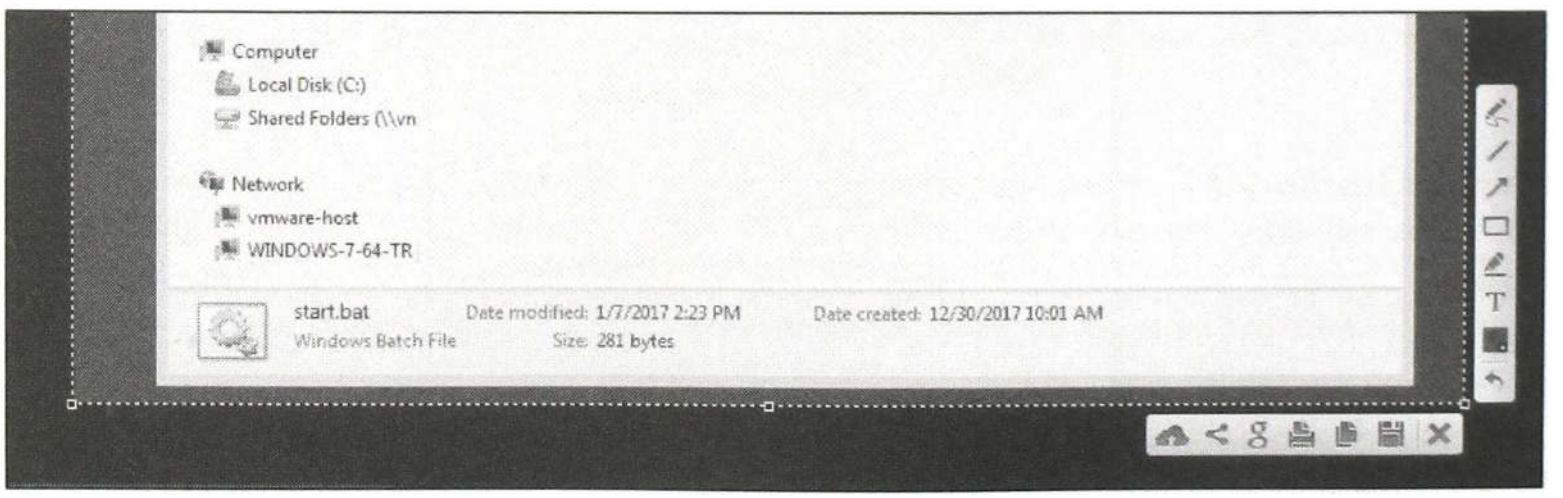


Figure 19.09: A Lightshot Capture application box and menu.

SmartDeblur (github.com/Y-Vladimir/SmartDeblur)

You may eventually locate a digital image of interest to your investigation that is blurry, distorted, or otherwise flawed. This happens to me often in the form of surveillance photos and video screenshots. There are commercial applications available that will assist with this situation, but they can be expensive. Another solution, Photoshop, is not only expensive, but difficult to use. My current free solution for this is SmartDeblur. Figure 19.10 (left) displays a blurry photo loaded within the SmartDeblur application. Adjusting the defect type, radius, smoothness, and correction strength will often clear up an image. Figure 19.10 (right) displays the same image after slight manipulation with the tool. The new settings are visible in Figure 19.11. Further adjustments will likely make the remaining text legible. I recommend a lot of practice with this program before it is needed for an image of importance.

Real World Application: I recently used this application to identify the license plate of a suspect vehicle as captured from a home surveillance system. The original image was too blurry to be helpful. The manipulated image provided all of the digits of the vehicle's registration. This technique is also used by several people on Reddit's "PicRequest" as explained in Chapter Seven.

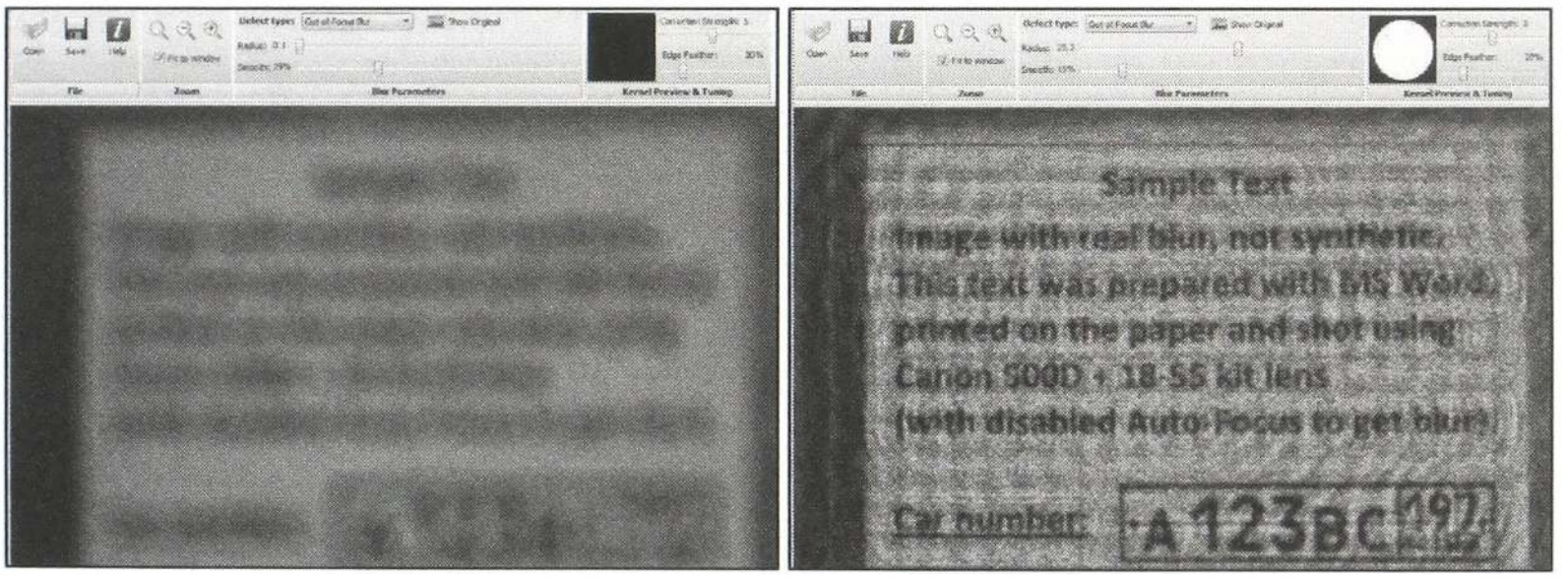


Figure 19.10: A blurry image opened in SmartDeblur.

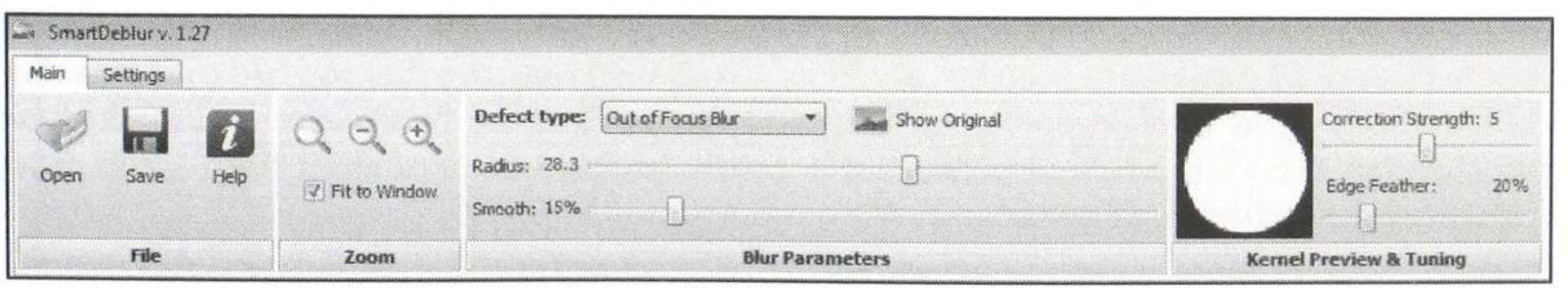


Figure 19.11: A SmartDeblur settings menu.

FOCA (github.com/ElevenPaths/FOCA)

FOCA has many uses. Some are legit, and some are a little sneaky. Everything that the program can do is legal, but I will only focus on the areas that should pass any ethics debate. FOCA's biggest strength is the ability to extract metadata from documents. You can drag and drop a document into the program, and it will analyze the file's hidden data and present a summary. Though there are many programs that can do this, FOCA stands out by automating the search process by searching, downloading, and analyzing all documents on a web server with very little input. The entire process follows.

When you launch FOCA, you will see the main program with few action choices. Click on "Project" and then "New project". Create a project name and provide the target website if you have one. Choose a location to store any documents located and include any notes necessary. Click the "Create" button to create your project. You will be prompted to name your project file, which will be the target domain name by default. Choose the same location as you had chosen to save the documents. You are now ready to begin analysis.

If you have any locally saved documents to analyze, you can now "drag and drop" them into the program. When the file is visible in the program, right-click on it and choose "extract all metadata". The left column will now have the analyzed content ready. Clicking on the file name under the "Documents" section will display the full metadata of the document. This will often include dates and times associated with modifications of the file; user names of people that modified it; printers that have printed the file; revision history; email addresses of the owner; and software version information. Figure 19.12 (top) displays a partial result of a file summary that displays the company, computer user name, and email address of the target. This information is probably stored inside every document created on that computer.

If you do not have any individual documents to analyze, you can begin to look for them now. On the same main menu, you should see various check boxes in the upper right corner. Uncheck the box next to "Exalead". You can try a search with this enabled, but it tends to cause problems due to Exalead's API rules. Now click the "search all" button. The program will use Google and Bing to search for any documents on the target website that were indexed by the search engines. In my example, it located 107 documents on the very informative website irongeek.com. When it is finished searching, right click on any of the files located and click "Download All". This will save a copy of all documents found to the location on your computer that you had chosen earlier. This can take some time depending on the amount and size of the documents. Even though we have not analyzed any of the documents yet, simply having a copy of all of them could prove beneficial for further intelligence.

After the files have finished downloading, right-click on any of them and select "Extract All Metadata". This will extract each document's raw metadata. When complete, right-click any file and select "Analyze Metadata". This will analyze all of the extracted content and categorize the results by various topics. The left column will now display several new subfolders under

"Documents". This analysis may take some time if there are a large number of documents. The first section will identify the documents by file type. Figure 19.12 (bottom) identifies 107 documents on the website irongeek.com. This includes 51 PDFs, 52 PowerPoints, and 1 Open Office document. The "Users" summary identifies five user names associated with the account, including the website owner, Adrian Crenshaw.

There are several scenarios where this program can prove itself valuable. People that run illegal websites often use fraudulent information during the registration process. If a website has been identified as a target, this registration (Whois) information becomes useless. Often, these subjects use a real name for the login name for their computers or as part of the registration of software. When this data is captured within the documents, you now have a lead on the identity of the target. As with most open source intelligence, this data is user created and you cannot always assume that the information is correct. However, it can provide great intelligence for additional searches. Another great reason to use this tool is to make sure that you did not miss any documents during the analysis of a website. Most sites have several pages in folders and subfolders. This is one way to locate all documents that are publicly available.

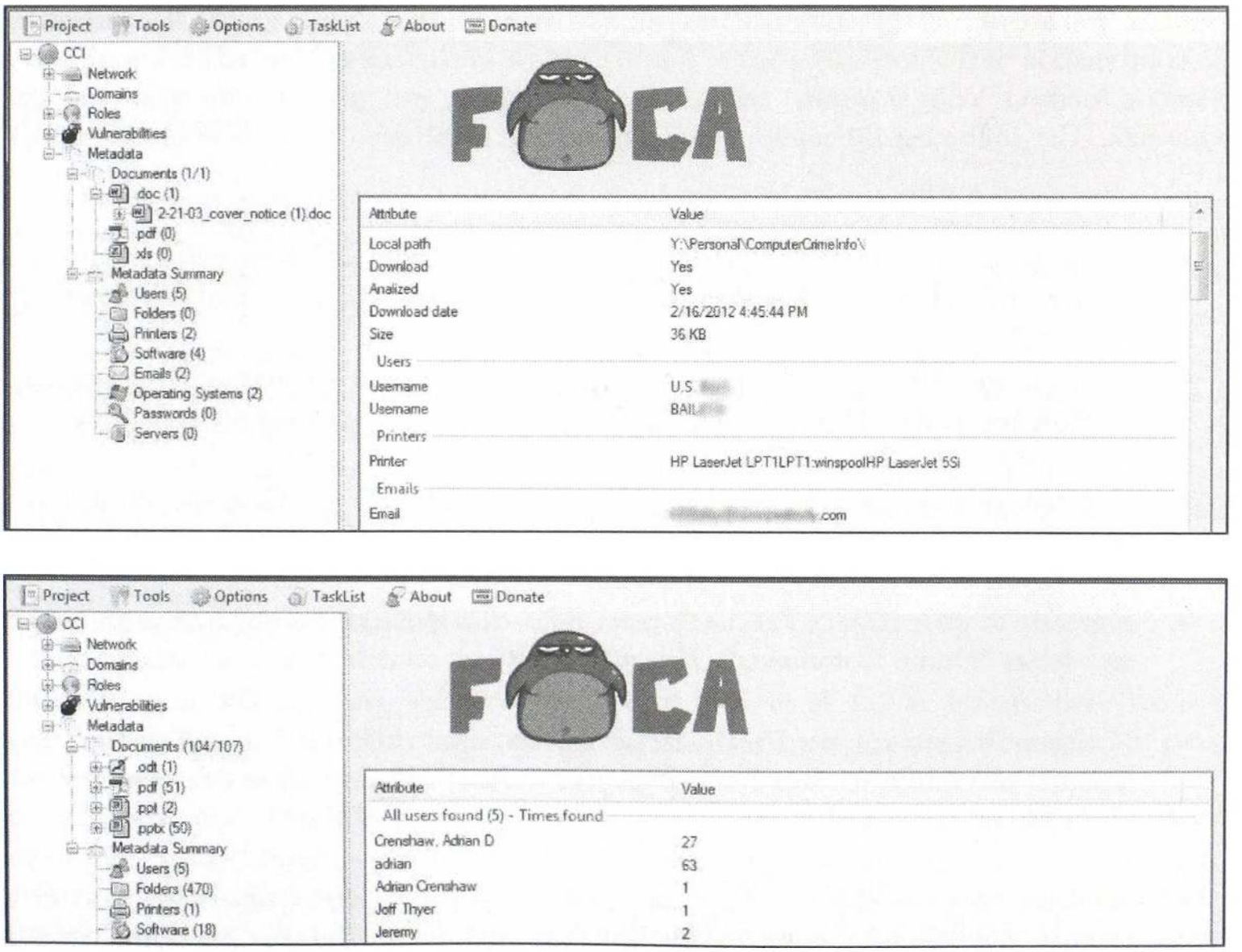


Figure 19.12: A partial FOCA analysis.

ExtractFace (le-tools.com/ExtractFace.html)

Investigators, researchers, or other analysts often need to get local copies of Facebook data. Facebook's interface has not been designed for that and does not provide printing or saving functions. If a user wants a copy of a conversation in his own profile, there is no simple solution. This program was designed to collect data from Facebook by providing many useful functions to automate tasks. It does not require installation, but does require the MozRepl Firefox add-on (addons.mozilla.org/en-US/firefox/addon/mozrepl). Additionally, at the time of this writing, it only worked with versions of Firefox prior to 57. Therefore, you will need Firefox Extended Support Release (ESR) available at https://www.mozilla.org/en-US/firefox/organizations/all/. Be sure to make this the default browser. This may seem like a lot of hassle in order to use one tool, but the rewards are worth the configuration.

I recommend selecting the "Activate on Startup" option under Tools > MozRepl in Firefox. In order to start using ExtractFace, run the executable file which will load the app in the taskbar menu. Right-clicking the icon in the taskbar presents the menu with options to expand a profile and "dump" a person's photo albums, friends list, and chat messages. When the collection is complete, you are prompted to choose a location to save a PDF document or XLSX spreadsheet. The combination of this software and the Firefox add-on creates an automated Facebook content collection method. Your computer will scroll, load, archive, and generate files while you watch hands-free. The following instructions will explain the most used features available in Figure 19.13.

- Navigate to your target's Facebook profile, right-click the ExtractFace icon in the taskbar, and select "Scroll and Expand". This will expand all posts and comments.

- Navigate to your target's Facebook Friends list, right-click the ExtractFace icon in the taskbar, and select "Dump Friends". This will save all friends into a spreadsheet.

- Navigate to your target's Facebook Photos list, right-click the ExtractFace icon in the taskbar, and select "Dump Albums". This will archive all photos into a folder.

- Navigate to your target's Facebook post, right-click the ExtractFace icon in the taskbar, and select "Dump Comments". This will archive all comments into a spreadsheet.

- Navigate to your target's Facebook Group list, right-click the ExtractFace icon in the taskbar, and select "Dump Group Members". This will save all group members into a spreadsheet.

- Navigate to your target's Facebook profile, right-click the ExtractFace icon in the taskbar, and select "Current Profile ID". This will display the user ID of the target for later use.

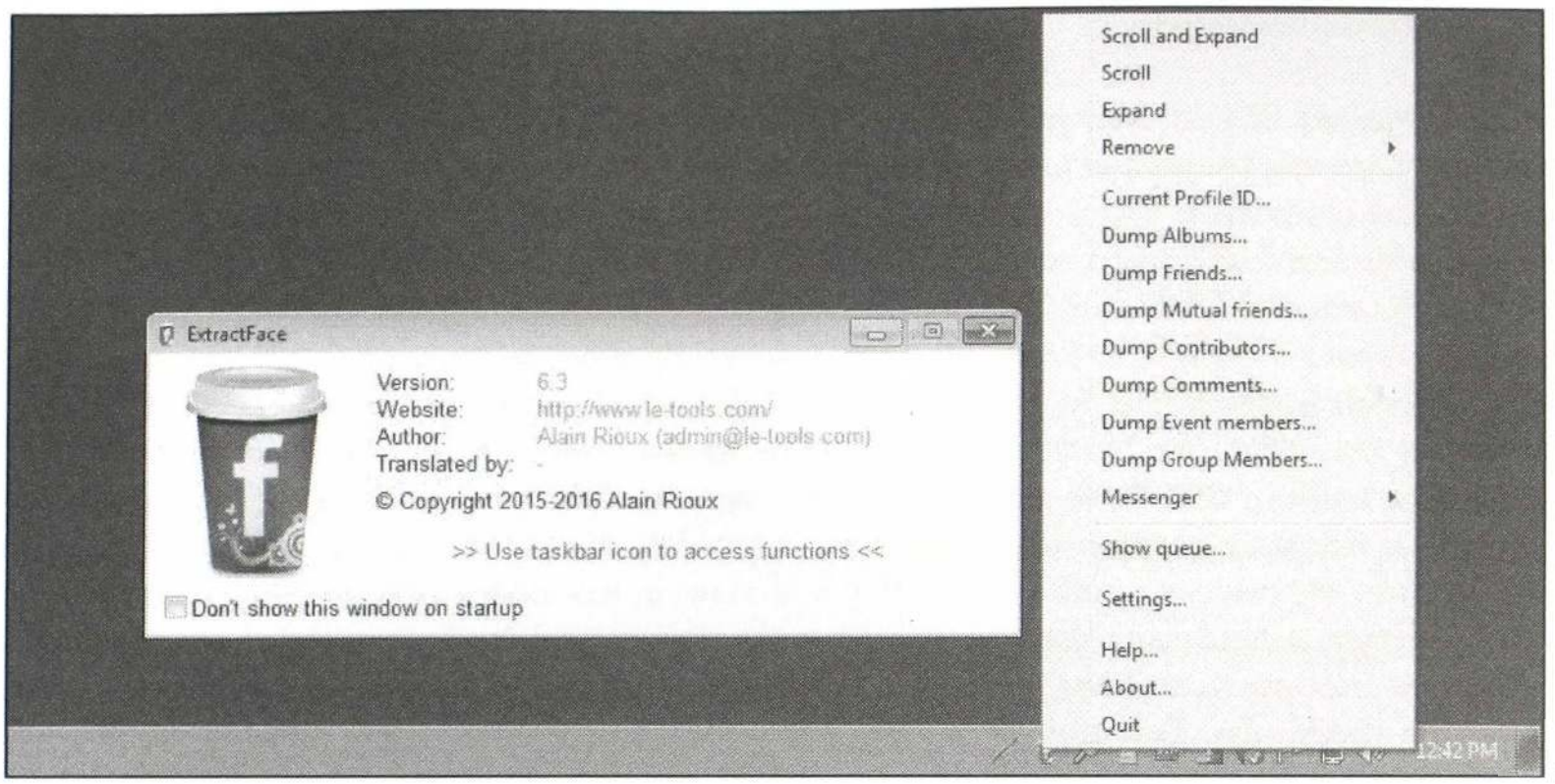


Figure 19.13: The ExtractFace menu (right) and "about" notice (left).

SEO Spider (screamingfrog.co.uk)

This program was designed to assist in Search Engine Optimization (SEO). This is the process of tweaking a website so that it will appear higher in search engine results. We can use it to identify all public pages, images, links, CSS files, and other data within our target's website. After installing and executing the free software, I conducted a search of "phonelosers.org". The entire process completed in approximately four minutes. Overall, this program identified 2,436 unique pages, 4,604 images, and 6,123 external links that were located within the target website. This tool will often identify much more data than a person can locate manually. On several occasions, this application has provided evidence that otherwise would have been missed. You can right-click on any result and open the data within a web browser. It will connect to the selected content within the target's live website. A complete user guide can be found at the above website.

Domain Hosting View (nirsoft.net)

This application will take any website domain and retrieve all public information about the registration and host. This will commonly include the names and contact information of people associated with the account. It will identify the physical host, which will be needed in the case of serving legal process. This is the same information as the "Whois" searches conducted in Chapter Sixteen but with a cleaner interface. You can only search one domain at a time with this tool. If you have several domains to search, you should use the WhoisThisDomain tool from the same website. Upon execution, it will present a window that will accept multiple IP addresses or domain names. The "View" menu offers a single combined report of all queries.

IP Net Info (nirsoft.net)

If you have an IP address instead of a domain, IP Net Info will perform the same series of searches for you. The first screen will allow you to enter multiple IP addresses and search them all. This is convenient when an investigator has received login IP addresses from an online provider such as Yahoo or Facebook. Both Domain Hosting View and IP Net Info will export the results as a text file for archiving.

Real World Application: In a child pornography investigation, I was supplied over 200 IP addresses that were used by the suspect while he traded illegal images of nude children. While I needed to confirm that these addresses were assigned to the suspect from his internet service provider, I first wanted to isolate the addresses according to the services that owned them. I used IP Net Info to conduct a bulk search of all addresses in less than a minute. The results allowed me to immediately identify the addresses that were most likely to be associated with his home internet service provider. I was also able to quickly see a pattern of a single non-residential address appear in the results. This address belonged to a local coffee shop where he used the Wi-Fi to facilitate his crimes. Video surveillance from this location proved extremely valuable. It showed him as the only person present that was on a computer. While this could have been accomplished manually, the automated tool allowed immediate results.

CCleaner (piriform.com/ccleaner/download)

This Windows and Mac utility is present on every machine that I use. In the simplest terms, it removes the junk left behind on your hard drive from daily usage. Internet cookies, cached files, registry errors, and many other issues can be resolved with this free software. The safest setting for both Windows and Mac users is to accept the default actions, "Analyze", review the files to be deleted, and then "Clean". Adventurous users may enable additional cleaning tasks. On my investigation machines, I enable everything in the operating system tab with the exception of user downloads and Custom Files and Folders, and everything in the applications tab. This would also wipe out any passwords stored in your browser, so be careful. The Registry section on Windows machines will eliminate unnecessary registry entries, while the Tools section on both Windows and Mac will help you control applications that automatically startup at boot. I like to disable anything that is not vital to my daily usage. I execute this program after every investigation.

BleachBit (bleachbit.org)

BleachBit is very similar to CCleaner, but can be a bit more aggressive. I select all available options with the exception of "Wipe Free Space". Choosing this would overwrite all free space on the hard drive which is time consuming. I currently only use this program on my investigative computers, including the Buscador virtual machine. Since I never operate these machines for personal use, I don't have concerns of deleting anything important. It should also be noted that I only execute CCleaner and BleachBit after I have archived any evidence obtained during the investigation.

VeraCrypt (veracrypt.codeplex.com)

VeraCrypt is an open-source utility used for on-the-fly encryption. It can create a virtual encrypted disk within a file which protects your evidence from unauthorized access. VeraCrypt is a fork of the discontinued TrueCrypt project. The idea is to create a place on your system where you can store all of your evidence, and encrypt it with a strong password. This would prevent accidental or intentional leakage, and can help justify the validity of your evidence in court. Imagine if you were asked during testimony about the number of people who could have potentially accessed the evidence in an investigation. With encryption, we can eliminate everyone but yourself.

- Download, install, and launch VeraCrypt. It is available on all platforms.

- Click Create Volume and the VeraCrypt Creation Wizard window should appear.

- Choose where you wish the VeraCrypt volume to be created. A VeraCrypt volume can reside in a file, which is also called a container, in a partition or drive. In this tutorial, we will choose the first option and create a VeraCrypt volume within a file. Click Next

- Choose whether to create a standard or hidden VeraCrypt volume. In this tutorial, we will choose the former option and create a standard VeraCrypt volume. Click Next.

- Specify where you wish the VeraCrypt volume (file container) to be created. Note that a VeraCrypt container is just like any normal file. It can be moved or deleted as any normal file. It also needs a filename, which you will choose in the next step. Click Select File.

- In this tutorial, we will create our VeraCrypt volume in the folder C:\Data\ and the filename of the volume (container) will be Evidence. Select the desired path in the file selector. Type the desired container file name (Evidence) in the Filename box. Click Save, then Next.

- Choose an encryption algorithm and a hash algorithm for the volume. If you are not sure what to select here, you can use the default settings and click Next.

- Specify the size of our VeraCrypt container. This needs to be large enough to store current and future evidence. Consider at least a few gigabytes. Click Next.

- Choose a strong volume password, and type it in the first input field. Then retype it in the input field below the first one and click Next.

- Move your mouse as randomly as possible within the Volume Creation Wizard window at least until the randomness indicator becomes green. The longer you move the mouse,

the better. This significantly increases the cryptographic strength of the encryption keys, which increases security.

- Click Format. Volume creation should begin. VeraCrypt will now create a file called Evidence in the folder C:\Data\. This file will be a VeraCrypt container. Depending on the size of the volume, the creation may take a long time.

- We have just successfully created a VeraCrypt volume (file container). In the VeraCrypt Volume Creation Wizard window, click Exit.

- In the remaining steps, we will mount the volume we just created. We will return to the main VeraCrypt window. Select a drive letter from the list (marked with a red rectangle). This will be the drive letter to which the VeraCrypt container will be mounted. Click Select File.

- In the file selector, browse to the container file (Evidence) and select it. Click Open.

- In the main VeraCrypt window, click Mount. Type the password.

- VeraCrypt will now mount the volume. We have just successfully mounted the container as a virtual disk. The virtual disk is entirely encrypted (including file names, allocation tables, free space, etc.) and behaves like a real disk. You can save (or copy, move, etc.) files to this virtual disk and they will be encrypted as they are being written. You can copy files (or folders) to and from the VeraCrypt volume just as you would copy them to any normal disk.

- If you want to close the volume and make files stored on it inaccessible, either restart your operating system or dismount the volume. Select the volume from the list of mounted volumes in the main VeraCrypt window and then click Dismount.

In this scenario, you would mount your encrypted evidence container before each investigation. As you locate evidence, you would save it within any desired folders created in the VeraCrypt container. These are now private only to you, and you can testify to the integrity of your evidence. While this may seem like a tedious process, it may halt any scrutiny toward the preservation of your evidence, and confirm to your opponents that you have gone above and beyond the requirements to protect your investigation.

KeePassXC (keepassxc.org/download)

Similar to VeraCrypt, KeePassXC creates a single file which is encrypted and protected with a strong password. The difference is that KeePassXC only protects your passwords, and does not store files. If you begin creating numerous covert online accounts, you will need a good password manager. This is my preference as it is easy to use, open-source, and free. Download, install, and execute KeePassXC, and then follow these directions to create and protect your first database.

- In the database menu select "New Database" and select the location to store it. Create a strong master password to protect the content.

- Click the "add new entry" button and provide the name of the website, your user name on the site, and a URL from where you would access the login page from. If you have a password already, enter it here. If not, consider allowing KeePassXC to generate a secure option for you.

- Click the icon of the black dice on the right side of the new entry window. Use the slider or the box at the end to specify the length of the password. When you are done generating a password, click on the "eye" to preview the password. Supply this password when you create or change a password to the website. Click "Apply" and "OK" to save the entry.

You have now saved a single password into your database. Repeat for all of your accounts as you create them. This program serves two main purposes. First, it securely stores your user names and passwords without connecting to the internet. This is a simple database. Second, it assists you with creating strong and random passwords that will better protect your accounts. As investigators, our online profiles may be attacked by those we are hunting. Make sure that your accounts are locked-down and immune to traditional password attacks.

Recuva (piriform.com/recuva/download/standard)

The final application to discuss in this chapter is Recuva. It scans your drive for deleted files and attempts to recover (Recu-va) any desired data. I include this because I have been contacted by numerous people that have accidentally deleted online evidence. I am guilty of this as well. Simply launch the software, select the drive that stored the files, and click Scan. The window below will populate with all of the files still present within the deleted area. After the scan completes, selecting the desired content and clicking the Recover button in the lower left quickly undeletes the files. This program has saved me many times, both during investigations and personal usage.

CHAPTER TWENTY
APPLICATION PROGRAMMING INTERFACES

This chapter will be the most technical of the book. It stands alone and is not required reading to understand the rest of the chapters. While none of these search methods are required to find OSINT information about a target, they will speed up the process and find new data. These techniques will automate the search process across several resources. The results will be text only pages, without ads or graphics, which eliminate confusing or unwanted content. If you are willing to invest a little time learning the process, the reward will be worth it.

An application programming interface (API) is an interface for software components to communicate with each other. APIs are used all over the internet. For example, when you use websites mentioned earlier such as AllMyTweets or Followerwonk, you are taking advantage of an API. These sites display Twitter information about a target. They obtain this information from the Twitter API. Twitter, just like many social sites, allows the public access to the API of its content. This allows developers to create new uses for the data, making it attractive to more users. Think of an API as a link to the large servers that hold all of the content of a website like Twitter. Going to twitter.com is one way to access the information stored by Twitter. Another option is to bypass the website and communicate with the API that has a direct link to all of the data on the servers. This API access will be more efficient for certain types of data.

Websites that take advantage of a single service's API are great, but they are just the tip of the iceberg. New services collect data from multiple APIs and combine the results. These APIs allow access to any user's entire social network information. If you supply a user name or email address, these services will fetch all online accounts associated with that subject. This can provide an immediate list of profiles that are associated with a target. This can save many hours of researching the information on traditional websites.

Some of the APIs detailed here will only provide information that you could otherwise obtain on the official website. The benefit of this API access is that only the content related to your target is presented. You will not receive sponsored links, ads, or misleading text and graphics. If you have multiple targets to research, using the API instead of the web interface will be much faster. You can even automate the search and execute hundreds of queries all at once. Ideally, you can create your own web form of all of these techniques to use when needed, such as mine visible in Figure 20.01. This chapter will explain the entire process. Before attempting any API searches, I recommend installing both the Firefox web browser and the mJSONViewer add-on. Both of these are explained in Chapter One. This combination of software will present all of the results within a standard web page. This page can then be printed, saved, or captured as a screen shot, the same way that you would document a website. Using Internet Explorer to open most API web pages in this section will produce undesired results.

Most of these services are designed to handle large volumes of requests. If you have a long list of names, email addresses, or telephone numbers, APIs can produce results much faster than traditional website searches. To take advantage of this automated feature, you will need to understand how scripts and batch files work. This is outside the scope of this book. If your agency has a need for bulk queries, it is worth approaching a programmer to create a customized solution. Otherwise, a few explanations should help understand the manual search techniques.

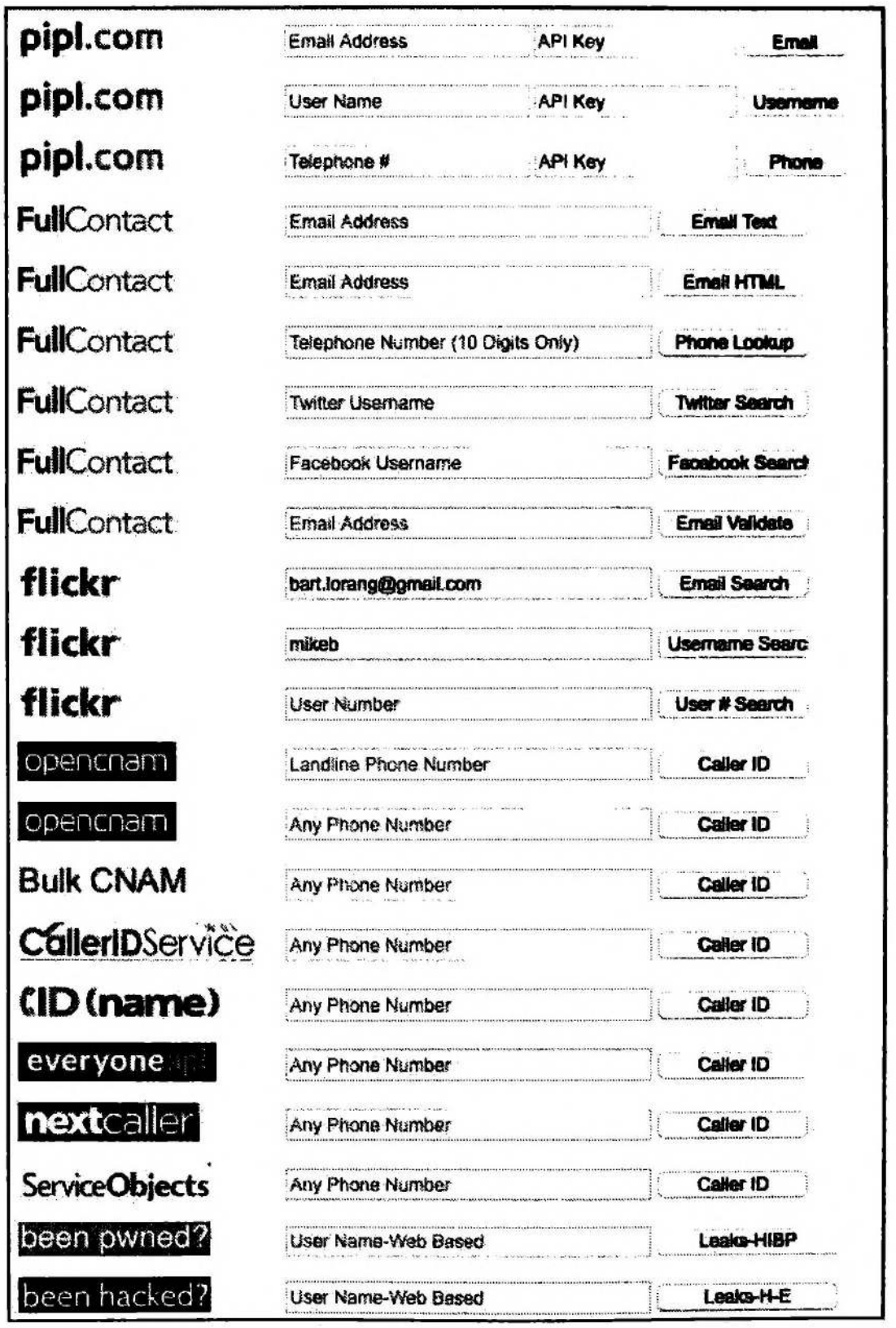


Figure 20.01: A custom API search page hosted at IntelTechniques.com.

Pipl (dev.pipl.com)

I explained earlier how Pipl can be a huge resource for information about a real name, email address, telephone number, or user name. The website is easy to navigate, but the API provides only the relevant data and it is easier to digest. A Pipl API key is required to conduct any searches and can be obtained for free at their site at pipl.com/api/demo. The XXX in all of these results should be replaced with your API key. The same API key will work for all four of the techniques detailed here. While Pipl does not limit the number of free trial API keys you can obtain, each is limited to ten queries.

Real name search:

https://api.pipl.com/search/?first_name=michael&last_name=bazzell&city=wood%20river&state=il&exact_name=false&no_sponsored=true&key=XXX

Email address search:

https://api.pipl.com/search/?email=michael@gmail.com&no_sponsored=true&key= XXX

User name search:

https://api.pipl.com/search/?user name=osintgeek&no_sponsored=true&key= XXX

Telephone number search:

https://api.pipl.com/search/?phone=##########&no_sponsored=true&key= XXX

These structured requests should start to look familiar now. In the previous examples, the data is detailed as follows.

https://api.pipl.com/search/: This tells Pipl to use the latest version.

first_name=michael&last_name=bazzell: This identifies the first and last name of the target.
city=wood%20river&state=il: This identifies the city and state of the target.
email=michael@gmail.com: This identifies the email address of the target.
user name=osintgeek: This identifies a user name of the target.
phone=##########: This represents an actual target telephone number.
&no_sponsored=true: This tells Pipl to exclude any advertisements.
&key= XXX: This represents your API key.

The most useful of these API requests is an email address search. When it is conducted, the associations of the target's email address with any social networks or online communities will be displayed. Any accounts created with the email address supplied might be included as a hyperlink.

These links could identify the personal websites of the subject. Clicking a link will forward you to the subject's site. You could now bookmark the URL that you created to get the information. Each time you wanted to check another email address, you could visit the bookmark and edit the email address to match that of your target. I prefer to create a form that simplifies this process. To do so, you need to create a web page. The page does not need to be uploaded anywhere on the internet, and you can execute the page from your computer's hard drive. If you have web development software, it will make it easy. If not, a text editor will work just fine. If you are a Windows user, open Notepad and type the following in a new text document.

```
<html><head></head><body><script type="text/javascript">
function dopipl1(firstpp, lastpp, citypp, statepp) {
window.open('https://api.pipl.com/search/v5/?first_name=' + firstpp + '&last_name=' +
lastpp + '&city=' + citypp + '&state=' + statepp +
'&exact_name=false&no_sponsored=true&key=XXX', 'pp1window');}</script>
<form onsubmit="dopipl1(this.firstpp.value, this.lastpp.value, this.citypp.value,
this.statepp.value); return false;">
<input type="text" name="firstpp" size="18" value="First Name" />
<input type="text" name="lastpp" size="18" value="Last Name" />
<input type="text" name="citypp" size="12" value="City" />
<input type="text" name="statepp" size="5" value="State" />
<input type="submit" /></form>
<script type="text/javascript">
function dopipl2(email) {
window.open('https://api.pipl.com/search/v5/?email=' + email +
'&no_sponsored=true&key= XXX, 'pipl2window');}</script>
<form onsubmit="dopipl2(this.pp2.value); return false;">
<input type="text" name="pp2" size="40" value="Email Address" />
<input type="submit" /></form>
<script type="text/javascript">
function dopipl3(ppuser) {
window.open('https://api.pipl.com/search/v5/?user name=' + ppuser +
'&no_sponsored=true&key= XXX, 'pipl3window');}</script>
<form onsubmit="dopipl3(this.pp3.value); return false;">
<input type="text" name="pp3" size="40" value="Screen Name" />
<input type="submit" /></form>
<script type="text/javascript">
function dopipl4(ppphone) {
window.open('https://api.pipl.com/search/v5/?phone=' + ppphone +
'&no_sponsored=true&key= XXX, 'pipl4window');}</script>
<form onsubmit="dopipl4(this.pp4.value); return false;">
<input type="text" name="pp4" size="30" value="Phone #" />
<input type="submit" /></form></body></html>
```

Save this file as pipl.html. When you open the file, it should appear in a web browser and look similar to the Pipl portion of Figure 20.01. A digital version of this file is maintained at https://inteltechniques.com/data/2018api.txt, which can be copied and pasted from the site. I maintain an online search tool with this code, at the following address.

https://inteltechniques.com/OSINT/api.html

Figure 20.02 displays this tool. Notice the first line requires an API key from Pipl. This only needs to be entered once, and you can populate queries on any of the options below. Figure 20.03 displays a Pipl API demo page with a valid key provided. Note that you do not need to provide any details in order to obtain this key. Simply visit the page at pipl.com/api/demo.

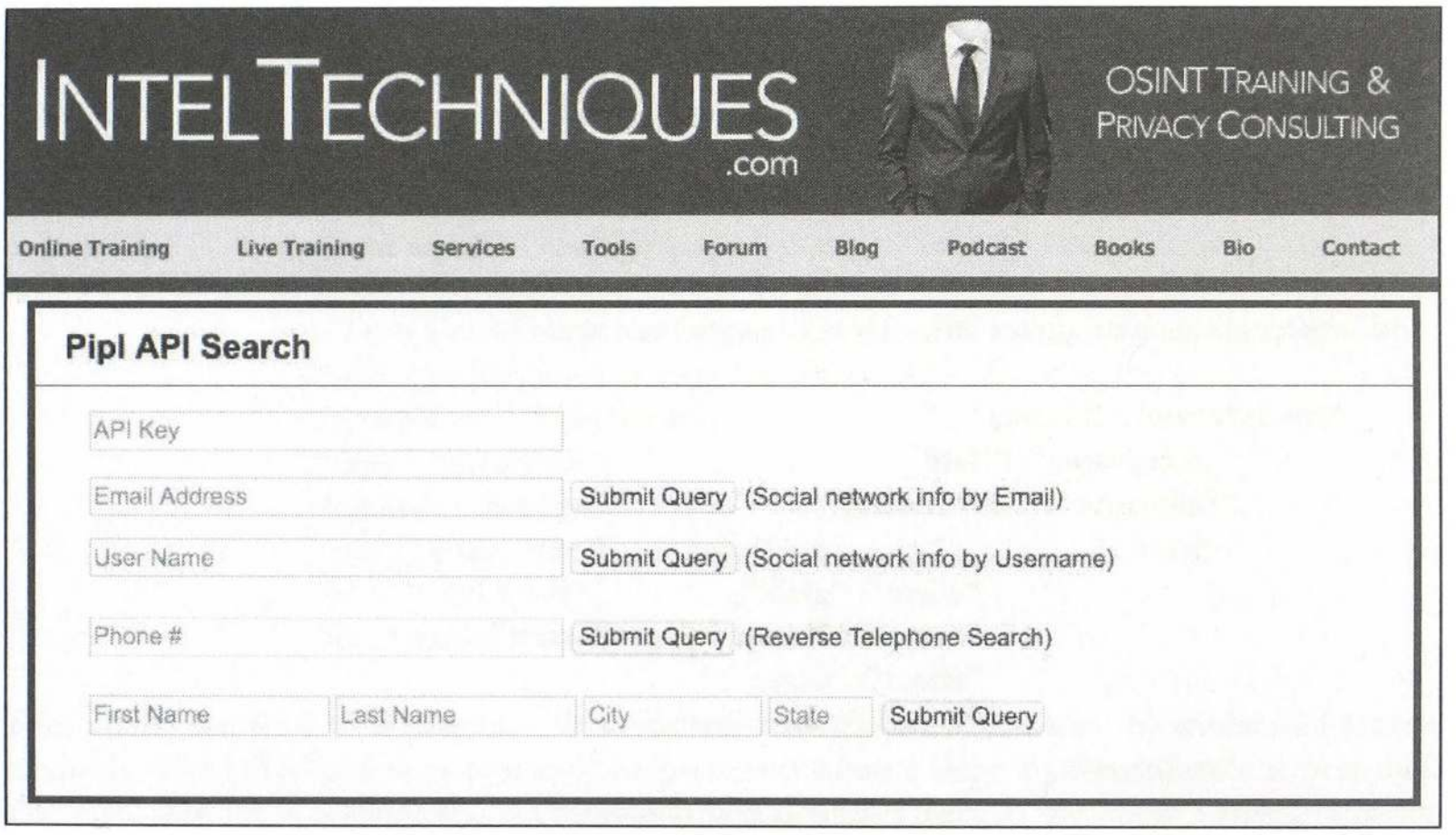


Figure 20.02: A Pipl API search page at IntelTechniques.com.

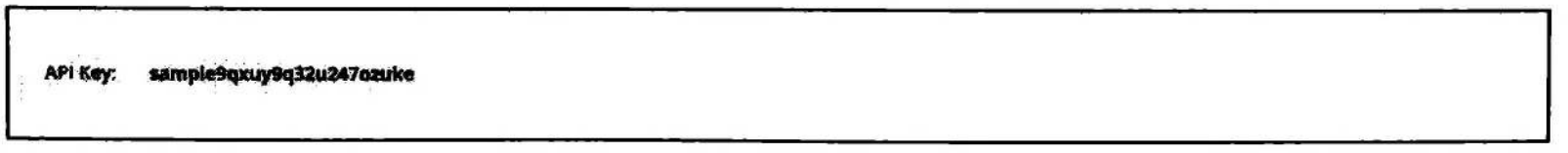


Figure 20.03: A Pipl API key generated at pipl.com/api/demo.

Full Contact (developer.fullcontact.com)

Full Contact provides a reverse search of any social networks or personal sites associated with the target email address. One of the advantages of this service is that it will attempt to find new information while you wait. A second search is occasionally required. Full Contact will scour the hundreds of social APIs for any relevant data. You can also specify in what format you want the results. This provides an additional option of an HTML view which provides icons and photographs. The following is the basic URL with a text view.

https://api.fullcontact.com/v2/person.json?email=lorangb@gmail.com&apiKey=XXX

This URL is detailed as follows.

https://api.fullcontact.com/v2/person: This refers to the site and version.
.json: This is the format output. Another option is HTML.
email=lorangb@gmail.com: This specifies the email address to search.
apiKey=XXX: This represents the API key which you will provide.

The result of this search includes dozens of associated social networks, photographs, profiles, and organizations with direct links. Next is a small excerpt.


```
"familyName": "Lorang",
        "givenName": "Bart",
        "fullName": "Bart Lorang",
        "chats": [
                "client": "gtalk",
                "handle": "lorangb@gmail.com"
                "client": "skype",
                "handle": "bart.lorang"
        "websites": [
                "url": "http://rainmaker.cc/"
"photos": [
        "typeName": "Twitter",
        "url":
        "http://a0.twimg.com/profile_images/1364842224/Bart_Profile_1_norm
        al.jpg"
    },
        "typeName": "Tungle Me",
        "type": "tungleme",
        "typeId": "tungleme",
        "url": "https://tungle.me/public/bartlorang/Image"
        "typeName": "Myspace",
        "type": "Myspace",
```

```
                "typeId": "Myspace",
                "url": "http://a2.ec-
                images.Myspacecdn.com/profile01/114/97c130815ed44e47a19080f97070
                6dbe/s.jpg"
],
"demographics": {
        "age": "32",
        "locationGeneral": "Denver, Colorado, United States",
        "gender": "Male",
        "ageRange": "25-34"
    },
"socialProfiles": [
        {
                "id": 5998422,
                "typeName": "Twitter",
                "following": 489,
                "followers": 662,
                "user name": "lorangb",
                "bio": "CEO & Co-Founder of @FullContactApp and @FullContactAPI
                - Tech Entrepreneur and Angel Investor.",
                "url": "http://www.twitter.com/lorangb",
                "typeName": "Facebook",
                "type": "facebook",
                "typeId": "facebook",
                "url": "http://facebook.com/bart.lorang",
                "id": "651620441",
                "user name": "bartlorang"
```


These links can lead to intelligence that may have taken hours to locate by traditional search methods. The HTML view option may be preferred when a large amount of data is available. The view displays a compressed result, which includes icon links to the target's social network profiles. The URL would be changed to the following address.

https://api.fullcontact.com/v2/person.html?email=lorangb@gmail.com&apiKey=XXX

In 2012, Full Contact added new search features to their API. You can now search by Facebook name, Twitter handle, and telephone number. Also, you can search an email address to identify the likelihood that it is an anonymous account. The Facebook search would appear as follows.

https://api.fullcontact.com/v2/person.html?facebookusername=bart.lorang&apiKey=XXX

Full Contact only requires the Facebook user name and not the entire address. In this example, "bart.lorang" is the user name, and the entire Facebook profile address would be

http://facebook.com/bart.lorang. This can be very useful when you do not know the email address of the target but have located a Facebook page. This would now help you identify all of the social networks of the target which should lead to the email address. The Twitter URL would appear as follows, if "bartlorang" was the Twitter handle.

https://api.fullcontact.com/v2/person.html?twitter=bartlorang&apiKey=XXX

The telephone search API of Full Contact is still in a beta stage, but I believe that this piece could prove to be very valuable in the future. If the target telephone number is associated with social networks, success is more likely. The search URL should appear like the following.

https://api.fullcontact.com/v2/person.html?phone=+16182590000&apiKey=XXX

The telephone number in this search is 618-259-0000. Neither hyphens nor spaces should be used when constructing the URL, and a "1" should precede the ten-digit number. The final offering from Full Contact provides a unique service that I have not encountered before. When you provide an email address, Full Contact offers information that may identify it as an anonymous account. If your target's email address is John@hushmail.com, most analysts would identify that domain as one that is associated with private email services. However, john@sharklasers.com may not seem so obvious. Both accounts belong to a service that provides disposable email accounts. Full Contact will attempt to identify these for you. The URL of the request should appear as follows, if the email address is john@shark.com. The result appears below the URL.

https://api.fullcontact.com/v2/email/disposable.html?email=john@shark.com&apiKey=XXX

```
<message>
Email user name is not sub addressed. Email's domain is likely associated with disposable email
addresses.
</message>
```

Creating a web form for all of these searches is done the same way as all of the rest. The following is the exact text that would create the page. Remember to change the XXX to your exact API key provided by Full Contact.

```
<html>
<head></head>
<body>
<script type="text/javascript">
function dofullcontact(femail) {
window.open('https://api.fullcontact.com/v2/person.html?email=' + femail +
'&apiKey=XXX', 'contactwindow');}</script>
<form onsubmit="dofullcontact(this.femail.value); return false;">
```

```
<input type="text" name="femail" size="40" value="Email Address" />
<input type="submit" /></form>
<script type="text/javascript">
function dofullcontact2(phone) {
window.open('https://api.fullcontact.com/v2/person.html?phone=+1' + phone +
'&apiKey=XXX', 'contactwindow2');}</script>
<form onsubmit="dofullcontact2(this.fcp.value); return false;">
<input type="text" name="fcp" size="40" value="Telephone Number (10 Digits Only)" />
<input type="submit" /></form>
<script type="text/javascript">
function dofullcontact3(twitter) {
window.open('https://api.fullcontact.com/v2/person.html?twitter=' + twitter +
'&apiKey=XXX, 'contactwindow3');}</script>
<form onsubmit="dofullcontact(this.fct.value); return false;">
<input type="text" name="fct" size="40" value="Twitter Handle" />
<input type="submit" /></form>
<script type="text/javascript">
function dofullcontact4(fb) {
window.open('https://api.fullcontact.com/v2/person.html?facebookusername=' + fb +
'&apiKey=XXX, 'contactwindow4');}</script>
<form onsubmit="dofullcontact(this.fb.value); return false;">
<input type="text" name="fb" size="40" value="Facebook Name" />
<input type="submit" /></form>
<script type="text/javascript">
function dofullcontact5(anon) {
window.open('https://api.fullcontact.com/v2/email/disposable.html?email=' + anon +
'&apiKey=XXX, 'contactwindow'5);}</script>
<form onsubmit="dofullcontact(this.anon.value); return false;">
<input type="text" name="anon" size="40" value="Email Address" />
<input type="submit" />
</form>
</body>
</html>
```

Save this file as fullcontact.html. When you open the file, it should appear in a web browser and look similar to the Full Contact portion of Figure 20.01. A digital version of this file is at https://inteltechniques.com/data/2018api.txt, which can be copied and pasted from the site.

Flickr

Flickr is a very popular photo sharing service owned by Yahoo. There are three specific uses of the Flickr API that I have found helpful during many online investigations. The first queries an email address and identifies any Flickr accounts associated with it. The second queries a user name, and identifies the Flickr user number of the connected account. The final option queries a Flickr user number and identifies the attached user name. The following html code could be copied into a text file, which would create your own search option similar to that in Figure 20.01. This text is available in the digital version located at inteltechniques.com/data/2018api.txt.

```
<script type="text/javascript">
function doflemail(flemail) {
window.open('https://api.flickr.com/services/rest/?method=flickr.people.findByEmail&api_k
ey=XXX&find_email=' + flemail, 'flickremailwindow');
}</script>
 <form onSubmit="doflemail(this.flemail.value); return false;">
<input name="flemail" size="40" placeholder="Email Address" type="text" />
<input type="submit" style="width:100px" value="Email Search"/>
</form>

<script type="text/javascript">
function dofluser(fluser) {
window.open('https://api.flickr.com/services/rest/?method=flickr.people.findByUsername&a
pi_key=XXX&username=' + fluser, 'flickruserwindow');
}</script>
<form onSubmit="dofluser(this.fluser.value); return false;">
<input name="fluser" size="40" placeholder="Username" type="text" />
<input type="submit" style="width:100px" value="Username Search"/>
</form>

<script type="text/javascript">
function doflnumber(flnumber) {
window.open('https://api.flickr.com/services/rest/?method=flickr.people.getInfo&api_key=X
XX&user_id=' + flnumber + '&format=rest', 'flickrnumberwindow');
}</script>
<form onSubmit="doflnumber(this.flnumber.value); return false;">
<input name="flnumber" size="40" placeholder="User Number" type="text" />
<input type="submit" style="width:100px" value="User # Search"/>
</form>
```

Reverse Caller ID Engines

Chapter Eleven explained the various reverse API methods that can identify landline and cellular telephone numbers. The API process was documented for searching each individual company for caller ID information. This section will explain how to create a web page search tool that will execute a search across all companies for a single telephone number. The following is the exact text that would create the page. Remember to change "XXX" on each line to your exact token key provided by each caller ID service.

```
<head></head><body>
<script type="text/javascript">function docidall(cidall) {
window.open('http://api.opencnam.com/v2/phone/+1' + cidall, 'frame1');
window.open('http://api.opencnam.com/v2/phone/+1' + cidall +
'?account_sid=XXX&auth_token=XXX', 'frame2');
window.open('http://cnam.bulkcnam.com/?id=XXX&did=' + cidall, 'frame3');
window.open('http://cnam.calleridservice.com/query?u=USERNAME&k=XXX&n=' +
cidall, 'frame4');
window.open('http://trial.serviceobjects.com/gppl/geophoneplus.asmx/GetPhoneInfo_V2?P
honeNumber=' + cidall + '&TestType=full&LicenseKey=XXX', 'frame5');
window.open('https://dip.cidname.com/' + cidall + '/XXX&output=raw&reply=none',
'frame6');}
window.open('https://XXX:XXX@api.nextcaller.com/v2/records/?phone=' + cidall +
'&format=json', frame7);</script>
<form onsubmit="docidall(this.cidall.value); return false;"><input type="text" name="cidall"
size="40" value="Phone Number" /><input type="submit" />(BASIC Caller ID Database)
<br /><br /></form>
<iframe allowScriptAccess='always' name='frame1' id='frame1' width='650px' height='40'
frameborder=0> </iframe>  <br />
<iframe allowScriptAccess='always' name='frame2' id='frame2' width='650px' height='40'
frameborder=0> </iframe><br />
<iframe allowScriptAccess='always' name='frame3' id='frame3' width='650px' height='40'
frameborder=0> </iframe><br />
<iframe allowScriptAccess='always' name='frame4' id='frame4' width='650px' height='40'
frameborder=0> </iframe><br />
<iframe allowScriptAccess='always' name='frame5' id='frame6' width='650px' height='40'
frameborder=0> </iframe><br />
<iframe allowScriptAccess='always' name='frame6' id='frame5' width='650px' height='540'
frameborder=0> </iframe>
<iframe allowScriptAccess='always' name='frame7' id='frame7' width='650px' height='540'
frameborder=0> </iframe></form></body></html>
```

Save this file as phone.html. When you open the file, it should appear in a web browser and look similar to Figure 20.01. A digital version of this file is at inteltechniques.com/data/api2018.txt.

Service Objects (serviceobjects.com/products/email/email-insight)

Service Objects offers several API search options based on sources of data. This company also offers an email lookup utility that will occasionally identify the location, age range, gender, income, education, and residence information of the user of the account. This is similar to TowerData, which will be discussed later. However, Service Objects displays more information during the free evaluation. Obtaining a free trial allows you up to 500 searches within a 30-day timeframe. Provide an email address, any name, and any telephone number to immediately be issued an active API key. Create a URL based on the following structure.

http://trial.serviceobjects.com/ei/emailinsight.asmx/GetContactInfoByEmail?Email=EMAIL&LicenseKey=XXX

I navigated to the following URL based on my API key and sample email address which produced the result seen directly below it.

http://trial.serviceobjects.com/ei/emailinsight.asmx/GetContactInfoByEmail?Email=m.wilson72@hotmail.com&LicenseKey=WS67-QRI1-OZH4

<City>Santa Barbara</City><County>Santa Barbara</County>
<State>CA</State><PostalCode>93105</PostalCode>
<AddressType>Residence</AddressType>
<Latitude>34.446925</Latitude><Longitude>-119.742822</Longitude>
<Age>45-54</Age><Gender>Male</Gender>
<HouseholdIncome>100K-125K</HouseholdIncome>
<Homeowner>True</Homeowner><HomeValue>350K-500K</HomeValue>

I now know that the data available for this email address indicates that the user is a male aged 45-54; he lives in Santa Barbara, California; and is a home owner with a residence valued at $350,000-$500,000. Creating a web page to search this data is very easy. Save the following code as serviceobjects.html. When you open the file, it should appear in a web browser and look similar to the other custom tools created in this chapter. A digital version of this file is at inteltechniques.com/data/2018api.txt, which can be copied and pasted from the site.

```
<html><head></head><body>
<script type="text/javascript">function doservice(email) {window.open
(' http://trial.serviceobjects.com/ei/emailinsight.asmx/GetContactInfoByEmail?Email=' +
email +'&LicenseKey=XXXX-XXXX-XXXX', servicewindow');}
</script>
<form onsubmit="doservice(this.service.value); return false;">
<input type="text" name="service" size="40" value="Email Address" />
<input type="submit" /></form>
</body></html>
```

TowerData (dashboard.towerdata.com/users/sign_up)

TowerData, formerly Rapleaf, is a company that builds products to analyze large amounts of information. They help businesses sort through large customer email databases and obtain information about the customers. They offer free access to the basic API that communicates with their data. This service will give us the location, sex, and age range of the user of a specific email address. It uses several sources of data including social networks and marketing data. For a small fee, you can obtain more information than these three categories, but this tutorial will only focus on the free data. TowerData's API, as well as most APIs, will require a unique API key. This is a license issued to you and no one else. Visiting the website above will allow you to create an account and request your key. Once you have the key, you need to construct a very specific URL, or internet address, to conduct the search, as follows.

https://api.towerdata.com/v5/td?email=test@test.com&api_key=xxx&format=html

This search is detailed as follows.

https://api.towerdata.com: This is the main domain of the service.
v5/td?: This identifies to TowerData that we are using the latest version (v5) of the API.
email=test@test.com: This specifies the email address that we want to search.
api_key=xxx: This represents your API key issued by TowerData.
format=html: This tells TowerData to present the results in a standard HTML view.

If you were to request this data using your API key and a valid email address, the result would be the following.

"location": "Chicago, Illinois, United States", "age": "25-34","gender": "Female"

You could now bookmark the URL that you created to get the information. Each time you wanted to check another email address, you could visit the bookmark and edit the email address to match that of your target. I prefer to create a form similar to those discussed previously:

```
<html><head></head><body>
<script type="text/javascript">function dotower(email) {window.open
('https://api.towerdata.com/v5/td?email=' + email +'&api_key=xxxx&format=html',
'towerwindow');}</script>
<form onsubmit="dotower(this.raf2.value); return false;">
<input type="text" name="raf2" size="40" value="Email Address" />
<input type="submit" /></form></body></html>
```

Save this file as towerdata.html. When you open the file, it should appear similar to Figure 20.01. You can now type in any email address, click Submit Query, and a new tab will open with the results from TowerData.

Compromised Email Addresses

In Chapter Eight, I explained how to query email addresses through services such as Have I Been Pwned and Hacked-Emails in order to identify database breaches that possessed the target account. I also explained in Chapter Nine how this same technique could be used to assume email addresses based on a user name. In that example, michaelb on Twitter could be associated with michaelb@yahoo.com, michaelb@gmail.com, etc. I demonstrated a custom online tool that allowed immediate query of these assumptions, which I will further explain here. You already know that you can navigate to hacked-emails.com, enter any email address, and receive a notification of any breaches that compromised that account. While this is a fairly quick process, the following URL will conduct the same query through the service's API.

hacked-emails.com/api?q=michael@gmail.com

You can conduct the same search on the Have I Been Pwned API at the following URL.

haveibeenpwned.com/api/v2/breachedaccount/michael@gmail.com?truncateResponse=true

The first query reveals the following partial response, with explanations in parentheses.

"title":"armyforceonline.com" (The breached website)
"author":"anon" (The hacker that posted the data)
"verified":false (Whether this is a "verified" leak)
"date_created":"2016-12-09" (The date the database was added)
"date_leaked":"2016-12-09" (The date the database was published)
"emails_count":1348214 (The number of email addresses in the leak)
"source_lines":1628872 (The number of lines in the database dump)
"source_size":90839257 (The size of the leaked database)
"source_network":"darknet" (The source of the leak)

As you can see, this provides much more detail than a simple notification that the target address was present in a specific leak. The address used in this demo provided several pages of results. Since these sites do not properly execute a user name search within assumed email addresses, we will create our own. The following code will provide a search form that will accept a user name as input. When executed, it will open several new tabs in your browser, each with a unique API search attempting to locate email accounts associated with your target user name. If you enter michaelb76, the results will include michaelb76@yahoo.com, michaelb76@gail.com, and several other popular email providers. As with the previous demonstrations, this entire text is available at https://inteltechniques.com/data/2018api.txt. Note that each result will populate within your browser approximately one second apart. This is to obey the rate limiting requirements of the API providers.

Have I Been Pwned

```
<script type="text/javascript">
function doall2(all2) {
setTimeout(function(){window.open('https://haveibeenpwned.com/api/v2/breachedaccount/
' + all2 + '@gmail.com?truncateResponse=true', '1leakwindow');},1000);
setTimeout(function(){window.open('https://haveibeenpwned.com/api/v2/breachedaccount/
' + all2 + '@yahoo.com?truncateResponse=true', '2leakwindow');},3000);
setTimeout(function(){window.open('https://haveibeenpwned.com/api/v2/breachedaccount/
' + all2 + '@hotmail.com?truncateResponse=true', '3leakwindow');},6000);
setTimeout(function(){window.open('https://haveibeenpwned.com/api/v2/breachedaccount/
' + all2 + '@protonmail.com?truncateResponse=true', '4leakwindow');},9000);
setTimeout(function(){window.open('https://haveibeenpwned.com/api/v2/breachedaccount/
' + all2 + '@live.com?truncateResponse=true', '5leakwindow');},12000);
setTimeout(function(){window.open('https://haveibeenpwned.com/api/v2/breachedaccount/
' + all2 + '@outlook.com?truncateResponse=true', '6leakwindow');},15000);
setTimeout(function(){window.open('https://haveibeenpwned.com/api/v2/breachedaccount/
' + all2 + '@icloud.com?truncateResponse=true', '7leakwindow');},18000);
setTimeout(function(){window.open('https://haveibeenpwned.com/api/v2/breachedaccount/
' + all2 + '@yandex.com?truncateResponse=true', '8leakwindow');},21000);
setTimeout(function(){window.open('https://haveibeenpwned.com/api/v2/breachedaccount/
' + all2 + '@gmx.com?truncateResponse=true', '9leakwindow');},23000);
setTimeout(function(){window.open('https://haveibeenpwned.com/api/v2/breachedaccount/
' + all2 + '@mail.com?truncateResponse=true', '10leakwindow');},26000);
setTimeout(function(){window.open('https://haveibeenpwned.com/api/v2/breachedaccount/
' + all2 + '@mac.com?truncateResponse=true', '11leakwindow');},29000);
setTimeout(function(){window.open('https://haveibeenpwned.com/api/v2/breachedaccount/
' + all2 + '@me.com?truncateResponse=true', '12leakwindow');},32000);
}

</script>
<form onsubmit="doall2(this.all2.value); return false;">
<input type="text" name="all2" id="ipleaks" size="40" placeholder="User Name-Web
Based" value="" />
<input type="submit" style="width:120px" value="Leaks-HIBP" />
</form>
```

Hacked-Emails

```
<script type="text/javascript">
function doall3(all3) {
setTimeout(function(){window.open('https://hacked-emails.com/api?q=' + all3 +
'@gmail.com', '1leakwindow');},1000);
setTimeout(function(){window.open('https://hacked-emails.com/api?q=' + all3 +
'@yahoo.com', '2leakwindow');},3000);
setTimeout(function(){window.open('https://hacked-emails.com/api?q=' + all3 +
'@hotmail.com', '3leakwindow');},6000);
setTimeout(function(){window.open('https://hacked-emails.com/api?q=' + all3 +
'@protonmail.com', '4leakwindow');},9000);
setTimeout(function(){window.open('https://hacked-emails.com/api?q=' + all3 + '@live.com',
'5leakwindow');},12000);
setTimeout(function(){window.open('https://hacked-emails.com/api?q=' + all3 +
'@outlook.com', '6leakwindow');},15000);
setTimeout(function(){window.open('https://hacked-emails.com/api?q=' + all3 +
'@icloud.com', '7leakwindow');},18000);
setTimeout(function(){window.open('https://hacked-emails.com/api?q=' + all3 +
'@yandex.com', '8leakwindow');},21000);
setTimeout(function(){window.open('https://hacked-emails.com/api?q=' + all3 +
'@gmx.com', '9leakwindow');},23000);
setTimeout(function(){window.open('https://hacked-emails.com/api?q=' + all3 +
'@mail.com', '10leakwindow');},26000);
setTimeout(function(){window.open('https://hacked-emails.com/api?q=' + all3 +
'@mac.com', '11leakwindow');},29000);
setTimeout(function(){window.open('https://hacked-emails.com/api?q=' + all3 + '@me.com',
'12leakwindow');},32000);
}

</script>
<form onsubmit="doall3(this.all3.value); return false;">
<input type="text" name="all3" id="ipleaks2" size="40" placeholder="User Name-Web
Based" value="" />
<input type="submit" style="width:120px" value="Leaks-H-E" />
</form>
```

Further Research

New APIs appear and disappear daily. Keeping up with the options can be exhausting. If this chapter has left you craving more detail, I recommend the following two services. I monitor both weekly for updates.

Programmable Web (programmableweb.com)

This website offers two extremely valuable services. The API search is a fairly standard option that allows you to query for specific services and identify any public APIs available. These could be for social networks such as Twitter or services such as Full Contact. Additionally, they offer a huge selection of "Mashups". These offerings are websites that combine two or more APIs and supply a public website for various queries. Many of the user name search services that I presented earlier were discovered through this option. New Mashups appear daily, and I browse them weekly.

Kong (konghq.com)

Kong is a collection of several APIs from several companies. They also offer their own API that can be used to access the data from other businesses. This collection includes both paid and free services, and most will allow you to test the product within this website. I am currently using this website for testing of individual email addresses through various Social Media Search APIs. Results can vary, but there is much potential with this type of combined services. I recommend creating a free account and experimenting with the various APIs.

As a closing reminder, I offer a pre-configured file of every API reference in this chapter. You can download the text file at https://inteltechniques.com/data/2018api.txt and replace each "XXX" with the API keys that you have obtained. You then need to change the file extension from txt to html in order for it to open with your web browser by default. Additionally, I host a web-ready version at the following address, but there are no API keys placed within the file. However, it will give you an indication of the look and feel of the custom form.

https://inteltechniques.com/data/2018api.html

Chapter Twenty-One
Android Emulation

For several years, online researchers have been navigating through various social networking websites for information about individuals. Whether it was older sites such as Friendster and MySpace, or newer networks such as Twitter and Facebook, we have always flocked to our web browsers to begin extracting data. Times have changed. Today, an entire generation of social network users rarely touch a traditional computer. They operate completely from a cellular telephone or tablet. Many of the networks through which individuals engage will only operate on a mobile device. Services such as SnapChat, Tinder, and Kik do not allow a user to access content from a traditional web browser. As this shift occurs, investigators must transition with it.

This chapter will focus on the huge amount of information available through mobile platforms that is not accessible through a web browser. I will explain a method of emulating a portable device within a traditional computer. Before we dive into the nuts and bolts of making things work, we should discuss why emulation is the way to go. In my investigations, documentation is my primary reason for launching a simulated mobile device within my computer operating system. If I conducted my investigation on an actual smartphone, documenting my findings can be difficult. Mobile screen captures only cover a small amount of visible content. Extracting any captured images can be a hassle. Referencing my findings within a final report can become very tedious. When using Android emulation within my traditional computer, I can easily create numerous screen captures, record a video of my entire investigation, and paste my results directly into the report.

Privacy and security are also important reasons to consider emulation versus directly investigating from a portable device. I have seen many law enforcement investigators conduct a search or use an app directly from their personal or work phones. This opens that device to scrutiny and discovery. An attorney could rightfully request a copy of the investigator's phone in order to conduct an independent forensic analysis. That would make most people nervous. Additionally, if I encounter malicious software or a virus from my portable device, it could affect all future investigations using that hardware. Emulation will remedy both of these situations.

The idea of Android emulation is to recreate the mobile operating experience within an application on your computer. This application will execute in the same manner that your web browser, word processor, or email client would open. It will have the exact same appearance as if you were staring at a telephone or tablet. Any actions that you take within this emulated device will not affect anything else on your computer. Think of it as an encapsulated box, and nothing comes in or gets out. A great feature of emulation is that you can create unlimited virtual devices. You could have one for every investigation in order to prevent any contamination.

Some readers will question why I chose to explain Android emulation instead of iPhone. The most obvious reason is the number of options. I will explain software solution for recreating the Android environment on your computer. The iPhone simulator will only function on Apple computers and has very limited features. The Android techniques will work on any major operating system. Additionally, we can create Android virtual machines that possess all original functionality. The iPhone simulator will not connect to most applications and features. There are more options for Android emulation than what I present in this chapter. My goal is to focus on the most user-friendly and feature rich solutions that are available without cost. My overall emulator of choice is Genymotion.

Genymotion (genymotion.com/download)

This application-based solution is extremely easy to use. It works with Windows, Mac, and Linux operating systems. I will provide details for the Windows installation, but the principles apply across all platforms. The operation of virtual devices after installation is identical on all operating systems. First, you will need to install the application.

Navigate to the Resources area of the website and click on the "Fun Zone" link. Next, click on the "Download Genymotion Personal Edition" link. You will be required to create an account for the free service. This can be completed using anonymous information and any real email address to which you have access. This setup file will contain both the Genymotion application and a virtual machine application called VirtualBox. Accepting all default installation options will install all of the required files. When the setup process has completed, you will have a new icon on your desktop titled Genymotion.

Execute this application and note that a Google Nexus 10 virtual machine is pre-installed and ready for launch. Instead of accepting this default option, consider creating your own machine in order to learn the process for future investigations. I recommend deleting this machine by clicking the trash icon to the right of the title. Perform the following instructions in order to create your custom Android devices.

- Create a new device by clicking the "Add" Icon in the menu. Select either the "Custom Phone – 7.0.0 – API 24 – 768x1280" or "Custom Phone – 6.0.0 – API 23 – 768x1280" options. This will create a new emulator of the previous public Android release (6.0.0) or the current release (7.0.0) in a default vertical telephone view (768x1280). It will not have any branding from a specific manufacturer. You may have a more recent version as you replicate this.

- Click Next and consider renaming this device similar to Case 18-123. Click Next again and allow the device to be created on your machine.

- Launch the new device by clicking the "Start" icon. The machine will load in a new window which should appear similar to the screen of an Android telephone. Click "OK"

to any feature notifications. Figure 21.01 (left) displays the default view of the home screen of version 7.0.0.

- Navigate within the Android emulator by single clicking on icons and using the "Back" icon in the lower left that appears similar to a left facing arrow.

- Consider the following customizations to improve the look and feel of the device. Figure 21.01 (right) displays the view of the home screen after these configurations have been made.

  Drag any app icons up and drop them in the "Remove" option.
  Click the "Applications" icon (six dots within circle), and choose "Settings".
  Choose "Display", then "Sleep", and select "30 Minutes".
  Choose "Security", then "Screen Lock", and choose "None".
  Press and hold the main window, select Wallpaper, and change if desired.

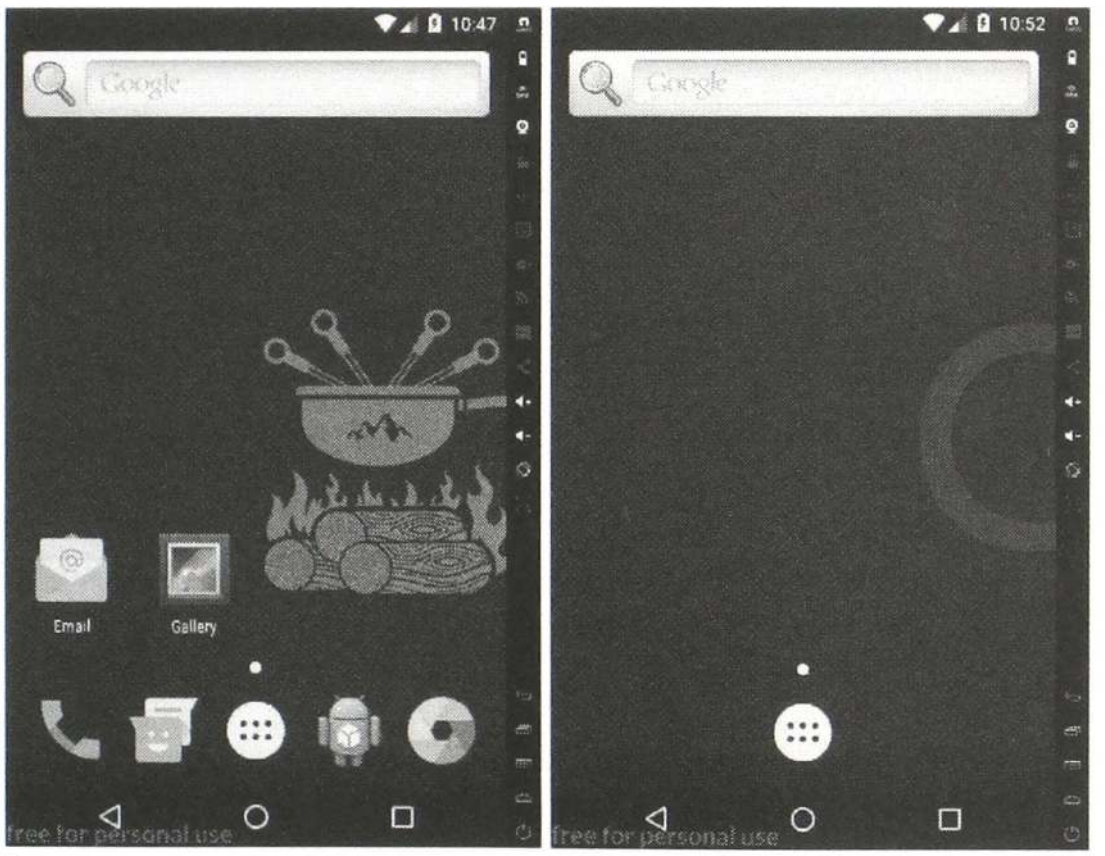


Figure 21.01: A default Android screen (left) and the custom version free of clutter (right).

You should now have a functioning replica of a standard Android device. However, you are missing several features. The biggest void is the absence of key applications such as Google Play and Gmail. Without core Google services, you cannot download apps to your device as part of your investigation tools. This has been the biggest hurdle with emulation. Consequently, there is finally an official fix, and an alternative option for advanced users. First, let's try the easy way by using the Genymotion built-in Google features.

- While inside our virtual Android device, click the "GAPPS" icon in the upper right corner. Accept the agreement and allow Google Apps to install. Select the option to restart the devices.

- Your browser will open to a specific page. Select "ARM", the version of the device that you created (7.0.0), and "Stock". Click the red download option in the lower right. Do NOT open the downloaded zip file.

- Drag-and-Drop the downloaded zip file into your virtual Android device. Accept any warnings. You may receive an error. When complete, close and restart the device.

You should now have the Google Play Store in your applications menu. Launching it should prompt you to connect to an existing or new Google account. Consider using an anonymous account that is not used for anything else. I do not recommend creating a new account from within this virtual machine because Google will likely demand a cellular telephone number for verification. I prefer to create the Google accounts from a traditional computer before connecting to the virtual Android device. After syncing with an active Google account on your new device, you should now be able to enter the Google Play Store. You should also now see all core Google services in your applications menu.

Many users have found this official way to fail them on occasion. I have found that many popular apps will not load with the most current versions of these utilities. Therefore, I offer an alternative option. Consider creating a second virtual device to have in case of future problems. The following instructions will restore the Play Store, emulate a more appropriate ARM driver (which will make some apps work better), and execute a patch that will eliminate those annoying Google crashes that have plagued this method for years.

- Click the "Add" button in the Genymotion menu and select "Custom Phone 6.0.0 – API 23". Allow the device to be created and launch it.

- Download the 6.0.0 ARM Driver at inteltechniques.com/android. Drag and drop the zip file directly into your running virtual Android device. Agree to the warning, and acknowledge the completion. Close the device and restart.

- Download the GApps 6.0.0 file at inteltechniques.com/android. Drag and drop the zip file directly into your running virtual Android device. Agree to the warning, and acknowledge the completion. Close the device and restart.

- Log in to a Google account that you will use to download apps to the device. Close any errors that appear. Close the device and restart.

- Download the Benzo Patch file at inteltechniques.com/android. Drag and drop the zip file directly into your running virtual Android device. Agree to the warning, and acknowledge the completion. Close the device and restart.

You should now have a fully-functioning Android 6 device with Google Play and no errors. You can now install any apps within the Play Store. If any apps refuse to install because of an incompatible device, download the desired app from **APK Pure** (apkpure.com) and drag and drop it into the machine. The addition of Google Play will allow you to natively install Android applications as if you were holding a real telephone or tablet. Launch Google Play and you will be able to search, install, and execute any app to your new virtual device. After you install a new program, click on the applications menu (circle with six dots in it). Click and hold the new app and you will be able to drag it to your home screen. Figure 21.02 displays the screen of my default investigation emulator. I will later explain how these programs can be used for intelligence collection. First, you should understand the features embedded into the Genymotion software.

When you launch an Android virtual machine, you will see a column on the right side of the window and a row of icons horizontally on the bottom. The bottom icons are part of the emulated Android system. Clicking the first icon will navigate you backward one screen from your current location. If you are within an app, this would take you back one step each time that you press it. The second icon represents the "Home" option and will always return you to the home screen. The third button is the "Recent Apps" option and it will load a view of recently opened applications.

The icons on the right of the emulator are features of Genymotion and allow you to control aspects of the Android machine from outside of the emulator. The following page displays this column of options, which should help explain each of these features. Note that many features are not available in the free version, but I have never found that to be a hindrance to my investigations. Genymotion is quite clear that if you plan on making money by designing an app through their product, you should pay for a license. Non-commercial usage allows unlimited use of the free personal version.

The GPS option within Genymotion is the most beneficial feature of their toolset. Clicking this icon will launch the GPS menu. Clicking the Off/On switch will execute the location spoofing service and a pre-configured GPS latitude and longitude will be provided. You can either supply the exact coordinates directly or click on the "Map" button to select a location via an interactive Google map. Figure 21.03 (left) displays the default GPS menu in the disabled state. Figure 21.03 (right) displays the menu with the exact coordinates of the Denver International Airport entered. I recommend changing the altitude, accuracy, and bearing settings to "0". Close this window and you will see a green check mark in the GPS button to confirm that your location settings are enabled.

GAPPS Indicator: Confirms Google Services are installed.

Battery Indicator: It does not have any impact on your virtual machine.

GPS: Enable and configure the current location reported to the virtual machine.

Webcam: Use your computer's webcam for live video within an app.

Screen Capture: Not available in the free version.

Remote Control: Not available in the free version.

Identifiers: Not available in the free version.

Disk I/O: Not available in the free version.

Network Configuration: Not available in the free version.

Phone: Not available in the free version.

App Sharing: Not available in the free version.



Volume Up

Volume Down

Screen Rotate: Flip your view into horizontal mode similar to a tablet.

Pixel Configuration: Not available in the free version.

Back Button: Moves back one screen from current app location.

Recent Apps: View recently opened applications.

Menu: Simulates the "Menu open" option within an application.

Home: Returns to the Home screen.

Power: Shuts down the device.

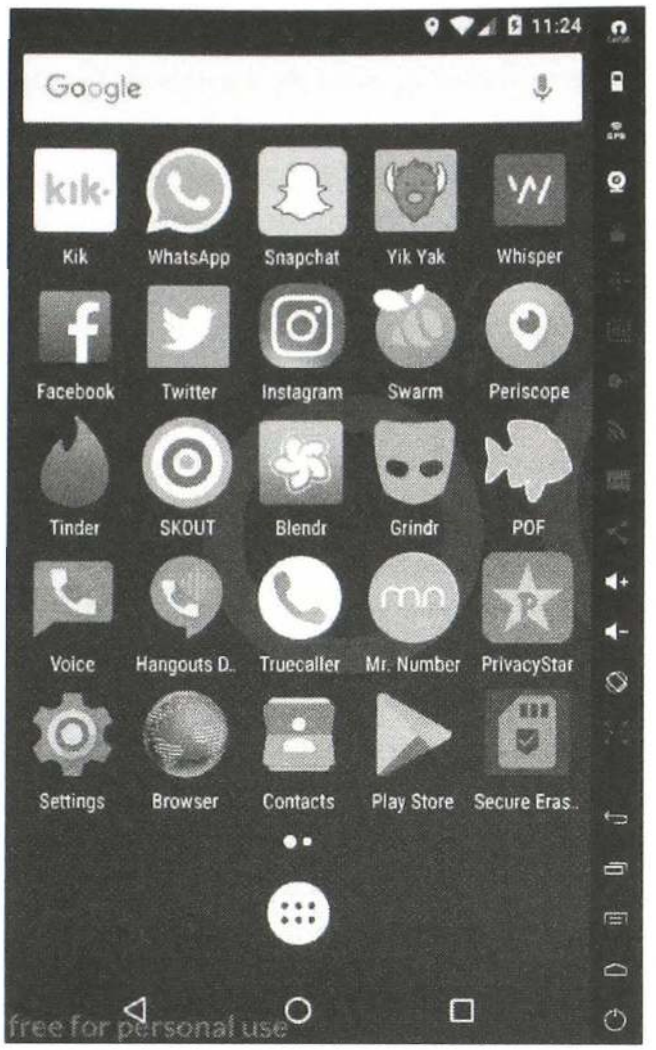


Figure 21.02: A custom Android emulator home screen with several apps installed into groups.

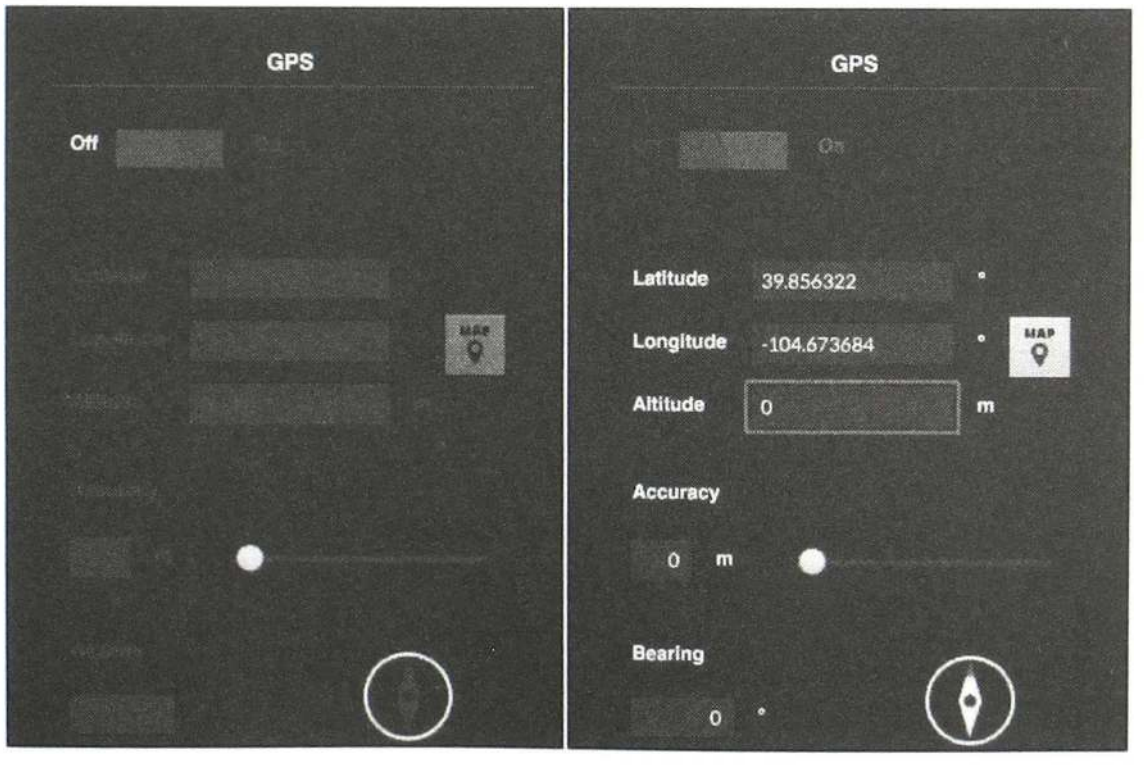


Figure 21.03: A disabled Genymotion GPS menu (left) and spoofed GPS (right).

Now that you have selected a location to broadcast through your device, you should test this configuration. My preferred way of doing this is to open Google Maps within the Android emulator and click the small blue target inside the white circle in the lower left. This will zoom Google Maps into the location where it believes you are located. You could also load the Bing Maps web page within a browser and ask it to center to your location. With both of these options, you may be prompted to "Allow" or "Deny" the device from obtaining your location details. You must choose "Allow" for this to function. After you have your desired location configured and you have confirmed accuracy, you can start to put this feature to work. The following tutorials will explain how location-aware applications could be used in investigations.

Facebook: The Facebook app on Android will appear similar to a compressed view of a standard profile page. The benefit of the mobile app is the ability to check into places. When you click the "Check In"" tab in the upper right corner, Facebook will present businesses near your current spoofed location. With my configuration, Facebook presented the terminals and airlines at the Denver airport. If you choose a location, and create a post on your timeline, Facebook will verify that you were there. I have used this when I need to portray that I am somewhere I am not. This method can help you establish credibility within your pseudo profile. You could easily create the illusion that you were working at a business all day or out clubbing all night.

Real World Application: I once helped a domestic violence victim confuse her ex-husband with this technique. I posted from her Facebook account "accidentally" leaving my spoofed location enabled. He stalked her every move online. After wasting his time going to random places trying to find her, and always finding the location to be closed, he began doubting the information that he uncovered about her whereabouts.

Twitter: The first time that you use Twitter within your Android environment, you will be asked if you want to share your location. While I usually discourage this type of activity, sharing your spoofed location can have many benefits. Similar to Facebook, you can make yourself appear to be somewhere that you are not. You may want to confuse your target. If you know that he or she will be monitoring your social networks using the techniques in Chapter Five, this method should throw them off and be misleading.

Yik Yak: This anonymous social media app is very popular with protesters and organized gatherings. It uses your location data and allows you to communicate with strangers within a ten-mile radius. With my configuration, I can watch people chat from the Denver Airport. I once used this app while investigating a group of looters that were ruining an otherwise peaceful protest. Because many of the undesirables participating in the looting do not know each other, they rely on apps like these to communicate. I was able to intercept live communication about the next business that they would target. When using, be sure to always expand the replies to individual posts. **Whisper** is another app that functions in nearly the same way.

Tinder: This dating app relies on your location in order to recommend people in your area that want to "hook up". It will use your Facebook account in use on your device for the login

credentials. The preferences menu will allow you to specify the gender, age range, and distance of the targeted individuals. Most people use this to identity members of their sexual preference within one mile of their current location. The users can then chat within the app. I have used this to identify whether a target was at home or another location. I have found that the most recent version of Tinder will not load on an emulated device. However, version 3.0.1 was working at the time of this writing. I keep a copy of the APK file at the following location.

https://inteltechniques.com/data/apk/tinder-3-0-1.apk.zip

Real World Application: During one investigation, I discovered that my target was a Tinder user. I set my GPS in my Android emulator to his residence. I could then search for men his age within one mile and identify if he was at home. If I did not get his profile as a result, I could change my GPS to his work address or favorite bar. When I received his profile in the results, I knew that he was near the spoofed location. I could do all of this from anywhere in the world.

Blendr/Badoo: These apps use the same database of user profiles. It is similar to Tinder, but does not require a Facebook account. This could be an additional option for locating a target that uses dating apps. The same method applied to Tinder would work on this network as well.

Skout: This app uses your Facebook account to populate data in your profile. The "Meet People" area will present individuals currently in the area of your supplied GPS coordinates. In addition to identifying the location of targeted individuals, this app could be used to identify people that are currently at a crime scene or gathering. I once used this technique to simply document people that were present near a state capitol during a credible bomb threat. When these people denied their presence during interviews, I had data that disagreed with their statements. Those that were lying quickly recanted their false statements and saved investigators a large amount of time.

Down: Formally called "Bang with Friends", this is another dating app based on your friends within your Facebook profile. If you have a covert profile that includes your targets as friends, launching this app will identify those friends that are also on Down.

Real World Application: I once used this during a cheating spouse investigation. I connected with a covert female Facebook profile that was friends with the suspected cheating spouse. Launching the Down app confirmed that he had an account. Swiping "Down" on his profile alerted him that I wanted to "get down" with him. This quickly resulted in a very incriminating chat that was later used in litigation.

FireChat: FireChat is a mobile app which uses wireless mesh networking to enable smartphones to connect via Bluetooth, Wi-Fi, or Apple's Multipeer Connectivity Framework without an internet connection. Though it wasn't designed with the purpose in mind, throughout 2014 it was used as a communication tool in civil protests. Launching the app within your emulator will identify live messaging and numerous chat rooms. These rooms can often contain valuable intelligence about live events from people directly involved. A social network login is not required

and people can use any pseudonym desired. I have used this only to obtain details of events. I have never had success in identifying individuals within the service.

There are many other similar apps. Now that you have an idea of how to integrate mobile applications into your investigations, you can apply the same techniques to the next future wave of popular apps. Many social network apps have no association with location. This content can still have value to an investigation. Some apps, such as Kik, only function within a portable device. You cannot load a web browser on a traditional computer and participate with these networks. However, you can access them from within your Android virtual machine. The following tutorials may help you find new uses for these popular apps.

Kik Messenger: Kik is an instant messaging application for mobile devices. It is modeled after BlackBerry's Messenger and uses a smartphone's data plan or Wi-Fi to transmit and receive messages. It also allows users to share photos, sketches, mobile webpages, and other content. You must create a free account within the app and you can then search any user name or Kik number. Many users do not share personal details, but you can still use the app during your investigation for covert communication with a target.

Real World Application: Child exploitation is prominent on Kik Messenger. Pedophiles have been quoted in news sources stating "I could go on it now and probably within 20 minutes have videos, pictures, everything else in between off the app. That's where all the child porn is coming off of" and "I can get anybody I want. I can achieve my sexual desires through this app". In 2014, a parent confiscated her 15-year-old daughter's cellular telephone after it was discovered that the minor was sending nude photos of herself to an older man at his request. I was able to use my Android emulator to log in as the child; continue conversations with the pedophile; and develop evidence to be used during prosecution. Documentation was easy with screen captures and screen recording.

WhatsApp: WhatsApp Messenger is an instant messaging app for smartphones that operates under a subscription business model. The proprietary, cross-platform app enables users of select feature phones to use the internet to communicate. In addition to text messaging, WhatsApp can be used to send images, videos, and audio media messages. Locations can also be shared through the use of integrated mapping features. It is the most globally popular messaging app with more than 800 million active users. You will need to create an account and provide a telephone number for verification. This number can be a cellular, landline, or VOIP number. I have had success using free Google Voice numbers. After you have an account, you can communicate directly with any target using the service. I have found that several of my targets refuse to converse over traditional text messaging, but freely text over WhatsApp. If you conduct any online covert operations, you should have this set up ahead of time.

SnapChat: As of 2018, I am no longer able to execute SnapChat on Genymotion. The latest versions block this behavior, and older versions require an update before launch. The only limited

success I have had is with Android Studio which exceeds the scope of this book. Even when I was able to force the program to function, constant errors prohibited actual use.

Text Messaging: If you conduct online investigations and communicate with a suspect, it is very possible that you may be asked to send or receive a standard SMS text message. Since your virtual device does not possess a cellular connection, and it is not assigned a telephone number, there are no native opportunities for this activity. However, you can install Google Voice and Hangouts Dialer within your virtual device. Voice will allow you to send and receive SMS text messages, and Hangouts Dialer will allow you to make and receive voice calls at the number issued by Google Voice. With this setup, you can conduct all of your communications over the virtual device, and preserve all of the evidence within a single archive.

Caller ID Apps: Chapter Eleven explained reverse caller ID services and how they can identify subscriber information associated with telephone numbers. There are several additional services that only support mobile use. Privacy Star is a powerful service that previously supported web search but now mandates that you install the app. After installation, you can search unlimited cellular and landline numbers in order to identify the owners. Other options include Mr. Number and True Caller.

Secure Eraser: As time passes, the size of your Android virtual devices will grow. System and app updates alone will increase the size of your files quickly. Much of this size is unnecessary. When these virtual machines download new data and update the files, the old files remain, and are not useable. Basically, your virtual devices start to take up a lot of space for no reason. Secure Eraser helps with this. On your master copy, after you have updated all of your software, launch Secure Eraser and change Random to 0000-0000. Click the start button and allow the process to complete. This will remove all of the deleted files. Restart your machine and then clone or export the device. The new copy will reflect the reduction of file size, but the master will still be large.

Contact Exploitation

Previous chapters identified ways to combine email accounts such as Gmail and Yahoo with social networks such as Facebook and Twitter in order to identify profiles connected to email addresses and cell numbers within your contacts list. Mobile apps are even more successful at retrieving profile information from your contacts. I have found that adding a cellular telephone number to the phone's address book will often obtain the following information relative to the target.

- Associated Facebook accounts from the "Find Friends" feature.
- Connected Google+ accounts from the Hangouts app.
- Google Play purchases and reviews from the Google Play Store.
- Associated Twitter accounts from the "Find Friends" feature.
- WhatsApp user names and numbers registered to the cell number.

Basically, entering a target's phone numbers and email addresses into your address book on an Android emulator forces many apps to believe that you are friends with the person. It overrides many authority protocols that would otherwise block you from seeing the connection from the real details to the connected profiles.

Virtual Device Cloning

There are several beneficial features that are disabled in the free version of Genymotion. Options such as cloning, resetting, sharing, and renaming are restricted unless you purchase a premium license. The following tutorials replicate these missing features, and do so legally. Genymotion relies heavily of VirtualBox to function. We can access our Android virtual devices within VirtualBox for better control of our investigations. Similar to the instruction in Chapter Two about using a clean virtual machine for every investigation, you should consider a new Android virtual device every time to research a target. The steps taken previously may seem too complicated to execute every day, so you may want to maintain a master copy and clone it.
The paid version of Genymotion will allow you to clone any machine. This is very beneficial when you have a custom emulator that contains a lot of configuration. You can instantly make a copy of that machine and use a new version for each investigation. However, the free edition of this software has disabled this feature. Instead, you will need to either manually create each machine that you want to use or clone the machine through VirtualBox. The benefit of manually installing and configuring your virtual devices is that you will keep your skills sharp. It should only take you ten minutes to create a new machine and incorporate the Google core services. However, that may be too time consuming if you use a new device for every investigation. The following instructions will clone the exact state of any virtual Android device within Genymotion.

- Create and customize an Android virtual device as desired. Configure Google Play and any other apps that you want present in all cloned copies. Optionally, execute the app "Secure Eraser" to eliminate unnecessary hard drive space. Exit the machine and close Genymotion completely.

- Open VirtualBox from your Applications folder (Mac) or Start menu (Windows). You should see the identical names of your Android machines visible within Genymotion. Right-click the machine that you want to duplicate and select "Clone". Figure 21.04 displays this program with a right-click menu option from an active machine.

- Provide a name for your new machine. This could be "Investigation Master Copy" or "2018-1234". Choose the options of Full Clone and Current machine state and click the Clone button. VirtualBox will create an exact duplicate of the chosen machine in the default folder for VirtualBox machines. You can identify this folder by right-clicking your new machine and choosing "Show in Finder" (Mac) or "Show on disk" (Windows).

You can now use this cloned device to conduct your investigation. Any changes made within it will have no impact on the master device. In fact, I title my master investigation devices "Master 7.0.0" and "Master 6.0.0". This way, I know to only open these to apply updates, and never for active investigations. Every time I need to use a device to research a target, I quickly clone the master and keep all of my cases isolated.

The presence of Android versions 6.0.0 and 7.0.0 may seem redundant. These numbers refer to the software version of the Android operating system within them. Possessing multiple versions is vital due to current and outdated apps. Most relevant apps will work fine on either version; however, some apps will not. I have encountered apps in the Play Store that would not allow installation under 6.0.0, but installed fine under 7.0.0. I have also witnessed the exact opposite. Having both versions available makes us prepared for either situation. You may be reading this book long after this writing and have more options present. Ultimately, I would have devices created for the most recent version and at least one previous version. If it is the year 2021, this may all be irrelevant.

Virtual Device Export

You may be asked to provide all digital evidence from your investigation as a matter of discovery. This could happen to a forensic examiner hired in a civil case or law enforcement prosecuting a criminal case. This is the precise reason that I create a new virtual device for all of my investigations. Not only is it a clean and fair environment, it is easy to archive and distribute when complete. The following instructions will generate a large single file that contains the entire virtual operating system and apps from your investigation.

- Exit Genymotion and open VirtualBox in the same manner as mentioned previously.

- Select the target virtual device, click on "File" in the menu bar, and select Export Appliance. Select the device again and provide the save location and name of the file. Figure 21.05 displays this menu and will save a file titled Android 5.1.0 2016.ova to the Documents folder of my Mac.

- Click Export and allow the process to complete. The final result will consist of a single file that can be archived to DVD or flash media.

- This file can be imported into VirtualBox by choosing the Import Appliance option in the File menu. This would allow another investigator to view the exact investigation environment as you.

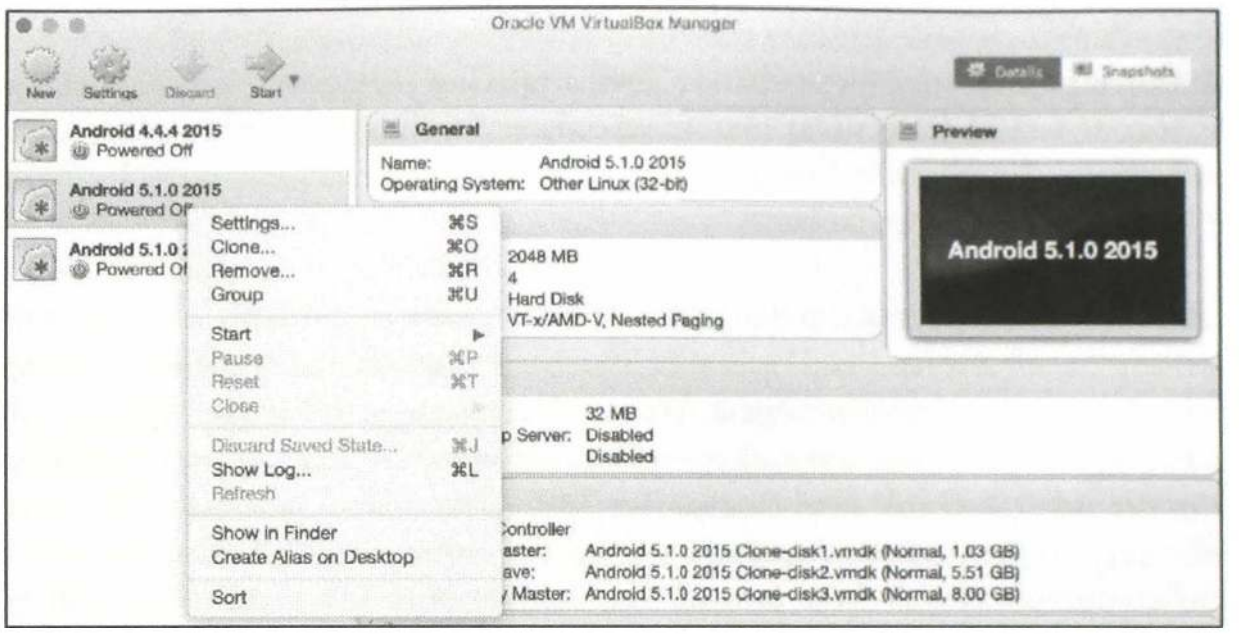


Figure 21.04: A VirtualBox menu with a clone option in the menu.

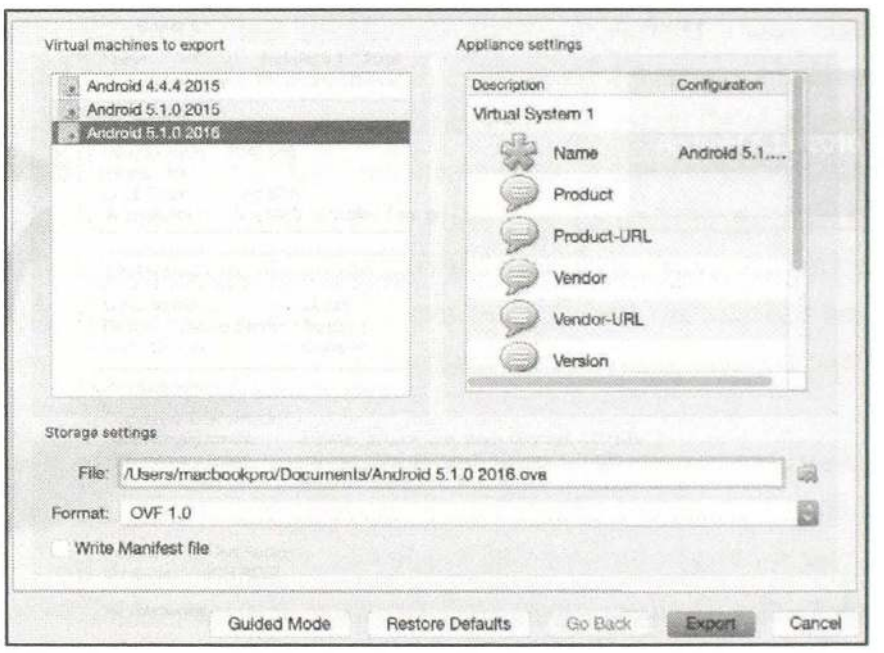


Figure 21.05: A VirtualBox export menu.

Genymotion is not your only option. **BlueStacks** (bluestacks.com), **Andy** (andyroid.net), and **NoxPlayer** (bignox.com) all offer the same basic functionality. After installation, most of these programs work the same way as Genymotion. I choose Genymotion over these because of the ability to import and export evidence as a virtual machine. While the others have their own backup and export options, I find the options presented here to be more transparent and acceptable in court. I encourage you to experiment with all of the options, and choose any that work best for you. Overall, I believe the future of OSINT collection will become more focused on mobile apps that have no website search option. In order to conduct thorough online investigations, mobile environment emulation is required. I highly recommend practicing these techniques with non-essential apps and data. This will better prepare you for an actual investigation with proper evidence control.

CHAPTER TWENTY-TWO
RECON-NG

Recon-ng is a full-featured web reconnaissance framework written in Python. Complete with independent modules, database interaction, built-in convenience functions, interactive help, and command completion, Recon-ng provides a powerful environment in which OSINT research can be conducted quickly and thoroughly. This utility provides automation to many of the redundant tasks that OSINT examiners find themselves performing on a daily basis. I offer a warning before proceeding. This is a technically complicated portion of this book. If you want to skip to the next chapter, you can always return when this type of application is needed. Recon-ng is already installed and configured in Buscador (Chapter Two), you only need to launch it from the menu. If you would like to install it within your own Linux environment, everything you need is at https://bitbucket.org/LaNMaSteR53/recon-ng/wiki/Usage%20Guide.

Recon-ng does not possess many online tutorials. The guides that I have found are mostly an index of commands with little explanation. Instead of trying to summarize how the program functions, I will walk you through actual usage and explain as we go. I will start with the basics and then conduct numerous actual searches. In lieu of screen captures, I will include all text input and output in **`9 pt Terminal bold font`**. Upon executing Recon-ng from the dock, a large portion of red text will appear followed by a default Terminal prompt. The red text indicates that you have not applied any API keys, and I will explain later how to do that. At this prompt, let's begin with the help command. Typing **`help`** reveals the following commands and explanations.

```
Add             Adds records to the database
Back            Exits the current context
Delete          Deletes records from the database
Exit            Exits the framework
help            Displays this menu
keys            Manages framework API keys
load            Loads specified module
pdb             Starts a Python Debugger session
query           Queries the database
record          Records commands to a resource file
reload          Reloads all modules
resource        Executes commands from a resource file
search          Searches available modules
set             Sets module options
shell           Executes shell commands
show            Shows various framework items
snapshots       Manages workspace snapshots
spool           Spools output to a file
unset           Unsets module options
use             Loads specified module
workspaces      Manages workspaces
```

Typing **show modules** will reveal the current functions available. Think of a module as a "resource". Just like Twitter is a website resource that we can use through a web browser, "twitter_mentions" is a specific resource that we can use in Recon-ng. The following modules were available at the time of this writing. We will use some of these during the instruction.

```
companies-contacts/bing_linkedin_cache
companies-contacts/jigsaw/point_usage
companies-contacts/jigsaw/purchase_contact
companies-contacts/jigsaw/search_contacts
companies-contacts/linkedin_auth
companies-multi/github_miner
companies-multi/whois_miner
contacts-contacts/mailtester
contacts-contacts/mangle
contacts-contacts/unmangle
contacts-credentials/hibp_breach
contacts-credentials/hibp_paste
contacts-domains/migrate_contacts
contacts-profiles/fullcontact
credentials-credentials/adobe
credentials-credentials/bozocrack
credentials-credentials/hashes_org
domains-contacts/metacrawler
domains-contacts/pgp_search
domains-contacts/whois_pocs
domains-credentials/pwnedlist/account_creds
domains-credentials/pwnedlist/api_usage
domains-credentials/pwnedlist/domain_creds
domains-credentials/pwnedlist/domain_ispwned
domains-credentials/pwnedlist/leak_lookup
domains-credentials/pwnedlist/leaks_dump
domains-domains/brute_suffix
domains-hosts/bing_domain_api
domains-hosts/bing_domain_web
domains-hosts/brute_hosts
domains-hosts/builtwith
domains-hosts/certificate_transparency
domains-hosts/google_site_api
domains-hosts/google_site_web
domains-hosts/hackertarget
domains-hosts/mx_spf_ip
domains-hosts/netcraft
domains-hosts/shodan_hostname
domains-hosts/ssl_san
domains-hosts/threatcrowd
domains-vulnerabilities/ghdb
domains-vulnerabilities/punkspider
domains-vulnerabilities/xssed
domains-vulnerabilities/xssposed
hosts-domains/migrate_hosts
hosts-hosts/bing_ip
hosts-hosts/freegeoip
hosts-hosts/ipinfodb
hosts-hosts/resolve
hosts-hosts/reverse_resolve
hosts-hosts/ssltools
hosts-locations/migrate_hosts
hosts-ports/shodan_ip
locations-locations/geocode
locations-locations/reverse_geocode
locations-pushpins/flickr
locations-pushpins/instagram
locations-pushpins/picasa
locations-pushpins/shodan
locations-pushpins/twitter
locations-pushpins/youtube
netblocks-companies/whois_orgs
netblocks-hosts/reverse_resolve
netblocks-hosts/shodan_net
netblocks-ports/census_2012
netblocks-ports/censysio
ports-hosts/migrate_ports
profiles-contacts/dev_diver
profiles-contacts/github_users
profiles-profiles/namechk
profiles-profiles/profiler
profiles-profiles/twitter_mentioned
profiles-profiles/twitter_mentions
profiles-repositories/github_repos
repositories-profiles/github_commits
repositories-vulnerabilities/gists_search
repositories-vulnerabilities/github_dorks
reporting/csv
reporting/html
reporting/json
reporting/list
reporting/proxifier
reporting/pushpin
reporting/xlsx
reporting/xml
```

Before we can conduct any research within this program, we must create a workspace. A workspace is a container that will isolate your work from one investigation to another. Think of a workspace as a case file. You may have a stack of cases on your desk, each with its own folder. All of your work on a case stays within the folder associated. Workspaces are similar. You should create a new workspace for each investigation. They can be deleted later or preserved for additional work. You can type **`workspaces list`** at any time to see the currently used workspaces. For now, we will create a new workspace titled OSINT by executing a command of **`workspaces add OSINT`**.

After creation, you will automatically begin using the new workspace. If you have created more than one workspace, such as one titled OSINT2, you can switch to it by typing **`workspaces select OSINT2`**. You might have a workspace for every target suspect or a single workspace for an entire case. Each situation will be unique. Now that you have a space created, we can begin.

The **`show`** command followed by a type of data will identify any stored content that will be used during an investigation. Typing **`show domains`** should reveal **`no data returned`** since we have not added any target domain names to our workspace. Therefore, let's add a domain name for our investigation by typing **`add domains social-engineer.org`**. Repeating the **`show domains`** command should now reveal the following.

```
+-----------------------------------------+
| rowid |        domain       |   module  |
+-----------------------------------------+
| 1     | social-engineer.org | user_defined |
+-----------------------------------------+
```

We have now added this domain into our system, and any module associated with domains that we execute will include this target. You could add every domain of interest and execute searches across all at once. Typing **`add domains cnn.com`** will make for a great example for the next option. Once you have your domains loaded, you may want to identify related web hosts. A domain, such as cnn.com, may have several unique host addresses that may be beneficial. Since I have already stored two domains in my system, typing **`use recon/domains-hosts/bing_domain_web`** will load a module ready to search across both. However, this command alone takes no action. After providing this command, you must type **`run`** and strike the enter key. This command checks the Bing search engine for hosts connected to the domains social-engineer.org and cnn.com. The result identified over 70 unique hosts, including the following.

```
[*] [host] internationaldesk.blogs.cnn.com (<blank>)
[*] [host] crossfire.blogs.cnn.com (<blank>)
[*] [host] reliablesources.blogs.cnn.com (<blank>)
[*] [host] lightyears.blogs.cnn.com (<blank>)
[*] [host] commercial.cnn.com (<blank>)
[*] [host] collection.cnn.com (<blank>)
```

We can replicate this type of search on Google to make sure that we are not missing any hosts that could be valuable by typing **use recon/domains-hosts/google_site_web**, striking the enter key, typing **run**, and striking the enter key again. This notifies us **35 total (15 new) hosts found**, which indicates that Bing found more hosts than Google, and Google found 15 hosts that we did not have in our collection from Bing. Since Recon-ng can parse out duplicates, we should have a list of unique hosts with a combined effort from both Google and Bing. Typing **show hosts** will display all of them.

We still have two domains stored in our workspace, but we may want to scan for additional options that we have not considered. There are many top-level domains (TLDs) aside from .com and .org. Executing **use recon/domains-domains/brute_suffix** and then **run** will scour the various TLDs such as .net, .tv, and others. After completion, typing **show domains** again will display our updated set of target addresses ready for further searching. In this example, I was notified that over 200 additional domains were located, mostly connected to cnn.com. These included numerous foreign versions and options such as cnn.org and cnn.photos. These are all new leads that should be analyzed later. We could now repeat our previous module execution of **use recon/domains-hosts/bing_domain_web** and likely grow our list of hosts substantially.

This is a good time to pause and consider what is happening here. As we find data, Recon-ng stores it and applies it to our searches. As those searches reveal more data, that content is added to our workspace. Every time we conduct a new search, or repeat a previous search, all of the stored data is applied, even the new content found recently. This prevents us from documenting everything that we locate because Recon-ng is keeping good notes for us. This can allow us to collect an amount of data otherwise impossible to manage manually. Let's move on to individual contacts.

Typing **show contacts** will display any contacts stored within the current workspace. You likely do not have any, so let's add some. Typing **use recon/domains-contacts/pgp_search** will use the recon module PGP search feature. It will scan all of the stored domains that we have located and search for any email addresses associated with public PGP keys within those domains. Typing **run** and striking enter executes the process, while submitting **show contacts** afterward displays the results. The following is the partial output with 33 new email addresses identified. Each of these addresses are now stored in your workspace, ready for the next round of research.

```
| rowid | first_name | middle_name | last_name |          email
+-------------------------------------------------------------------------
| 1     | Christopher |             | Hadnagy   | logan@social-engineer.org
| 2     | barsuk      |             |           | barsuk@cnn.com
| 3     | Tristan     |             | Helmich   | tristan.helmich@cnn.com
| 4     | Paul        | P           | Murphy    | paul.p.murphy@cnn.com
| 5     | Guy         |             | Incognito | Huzudra@cnn.com
| 6     | bob         |             |           | hello@cnn.com
| 7     | Jose        |             | Pagliery  | Jose.Pagliery@cnn.com
| 8     | D           | Ian         | Hopper    | ian.hopper@cnn.com
| 9     | Scott       | John        | Anderson  | scott.anderson@cnn.com
```

One of the most powerful email search options available is the Full Contact API, which was explained in Chapter Twenty. Hopefully, you have already obtained a free trial API key. If not, one can be requested at dashboard.fullcontact.com/register. In my examples, I will replace my actual Full Contact API key with XXX. We can load the Full Contact module by typing **use recon/contacts-profiles/fullcontact** and then **run**. You should receive an error message since you have not added any API keys to your copy of Recon-ng. At any time, you can type **keys list** and see any stored keys. Typing **keys add fullcontact_api XXX** will add your key for future use. Executing **run** again should now make the module function. Recon-ng is now searching all of the stored contacts through the Full Contact database. This will identify associated social networks. After completion, I was notified that 35 new profiles were added to my workspace and 8 new contacts were found. Basically, Recon-ng extracted all information it could from Full Contact and populated our database. The following is a small portion of the details available by typing **show contacts**.

```
Christopher Hadnagy (chris@social-engineer.com) Carnegie Mellon University
```

Below is a portion of the social network profiles added to our database. I obtained this by typing **show profiles** and striking enter.

```
-------------------------------------------------------------------------------
        username          |   resource   |                    url
-------------------------------------------------------------------------------
logan@social-engineer.org | Facebook     | https://www.facebook.com/chris.hadnagy
107828765414608142723     | GooglePlus   | https://plus.google.com/10714608142723
chrishadnagy              | Gravatar     | https://gravatar.com/chrishadnagy
christopherhadnagy        | LinkedIn     | linkedin.com/in/christopherhadnagy
```

Let's reflect on how this can be beneficial. Assume that you are investigating a website. Recon-ng has now identified people associated with the domain; email addresses connected to the people; and social network profiles created by the email accounts. Magnify this by tens or hundreds of subjects, and you have an easy way to replicate several hours of work. In another scenario, you are investigating a list of potential email addresses connected to a case. Entering these into Recon-ng allows you to execute your searches across all accounts. The effort to check one address is the same to check thousands. This impressive capability is only a small fraction of what can be done with this application. Now that we have a few profiles in our database, let's find more.

Typing **use recon/profiles-profiles/profiler** and striking enter loads the profiler option. Typing **run** executes the process which attempts to identify additional online services that possess accounts with the same user name as those in your database. In Chapter Nine, I explained manual services such as Know'em that display potential accounts of your target on other networks. This option is very similar, but queries any stored profiles within your workspace. After executing during this example, I typed **show profiles** which revealed the following partial output. This

action added several additional profiles of our target. Again, imagine how much time this could save if you had dozens of user names obtained through Recon-ng or added manually.

```
logan@social-engineer.org VideoLike videolike.org/video/logan@social-engineer.org
chrishadnagy              Klout     klout.com/chrishadnagy
chrishadnagy              VideoLike videolike.org/video/chrishadnagy
```

Now that we have several user names of targets in our workspace, we should consider searching within Twitter for any content of interest. We could navigate to Twitter, assume our suspect is chrishadnagy, and look for people mentioning him with to:chrishadnagy. Alternatively, we can ask Recon-ng to conduct this task for every user name we have collected. Typing **use recon/profiles-profiles/twitter_mentions** loads the module and typing **run** executes the process. You should immediately receive an error since you do not have a valid Twitter API key configured within Recon-ng. The process is identical to the Full Contact requirement, and the following commands will get you started. You can obtain your own Twitter API key and "secret" at apps.twitter.com.

```
keys add twitter_api xxx
keys add twitter_secret xxx
```

Repeating the **run** command executes the process, but nothing was found during this demonstration. Since I know that his Twitter user name is humanhacker, I can type the following to manually add that user name to our database. Note that you will be prompted for each response.

```
add profiles
username (TEXT): humanhacker
resource (TEXT): twitter
url (TEXT): https://twitter.com/humanhacker
category (TEXT): social
notes (TEXT): manual
```

Repeating the **run** command replicates the previous process, but now includes our new target name. The results identify the Twitter accounts that have mentioned our suspect at some point. Each of these have also been added to our growing list of profiles. Executing **show profiles** displays these new additions. If desired, you could repeat the **run** command again, and Recon-ng would attack these new user names, looking for other Twitter users that have mentioned them. This can quickly escalate to a point that you are identifying subjects with no interest to your case. This seems like a good time to back away, create a report, and start a new set of actions. The following commands will instruct Recon-ng that we want to use the reporting tool; mandate a graphical html (web) template be used; set the "Customer" as IntelTechniques; set the "Creator" as M.Bazzell; and execute the process.

```
use reporting/html
set CUSTOMER IntelTechniques
```

```
set CREATOR M.Bazzell
run
```

Note the output after the final command. It identifies that the report is complete, and provides the storage location. Since I am running Recon-ng from my Buscador virtual machine, the default location is /home/osint/.recon-ng/workspaces/OSINT/results.html. Therefore, I can open the home folder on my desktop; double-click the ".recon-ng" folder; double-click the "workspaces" folder; double-click the "OSINT" folder; and then open the "results" file. Figure 22.01 displays the partial file from this example. Note that the Domains, Hosts, and Contacts sections are not expanded, but contain a lot of information. At the bottom of this file, the "Created by", date, and time clearly identify these report details.

Hopefully this demonstration explained the usage of Recon-ng. Executing **exit** in the window closes everything, but removes nothing. Before our next example, let's delete our previous work and start fresh. Note that deleting a workspace removes all associated data and reports. Make sure that you have exported your evidence if needed. First, relaunch Recon-ng from the Buscador menu. The following commands display the current workspaces; delete the OSINT workspace; and create a new workspace titled location.

```
workspaces list
workspaces delete OSINT
workspaces add location
```

This demonstration explains automated options during an investigation of a location. Assume that you are researching a terrorism-related incident or crime scene. In these scenarios, you know the location, but do not know any people associated with the geographical area. I presented web-based services that individually allow location-based lookups on networks such as Twitter and Flickr. We can automate this with Recon-ng. Typing **add locations** will prompt you with **latitude (TEXT)**. Simply strike enter twice to bypass the latitude and longitude prompts (we don't know them yet). You will then be prompted with **street_address (TEXT)**. In this example, I will enter **1060 W Addison St, Chicago, IL 60613**. This stores the entered location into our current workspace.

The following commands load the geocode module which will convert our address to GPS coordinates; execute the module; display and confirm our locations; load the Twitter "pushpin" module; execute the script; and display the Twitter posts that were geo-tagged near our location.

```
use recon/locations-locations/geocode
run
show locations
use recon/locations-pushpins/twitter
run
show pushpins
```

We can replicate the same process for Flickr, but I must first add my Flickr API key, represented by "XXX". The following commands will add your Flickr API key; load the Flickr "pushpin" module; execute the script; and display the pushpins created.

```
keys add flickr_api XXX
use recon/locations-pushpins/flickr
run
show pushpins
```

The output of these commands included thousands of photos posted to Flickr that were geo-tagged within the area of the target location. If you encounter errors stating that your API keys are invalid after entering and running a script, simply exit Recon-ng and relaunch. This is a known bug. Now that we have several pushpins created, let's create the report. The following commands load the reporting module; display the stored location coordinates; set our latitude for the report; set our longitude for the report; define our radius as one kilometer; and execute the process.

```
use reporting/pushpin
show locations
set LATITUDE 41.9474536
set LONGITUDE -87.6561341
set RADIUS 1
run
```

Upon completion, your browser should launch two tabs. The first is a text listing of the Tweets and images located, and the second is an interactive map of the area. Clicking on any pushpin displays the associated details. Figures 22.02 and 22.03 display the pages created during this demonstration. While an entire book could be written about the possibilities with Recon-ng, let's conduct one last example to illustrate another way to use the program. If you have not already done so, type **exit** to close the window, then relaunch Recon-ng. The following commands will display your current workspaces; delete the previous example; and create a new space titled email.

```
workspaces list
workspaces delete location
workspaces add email
```

In this example, we want to manually add a contact, and you may want to add several targets. Typing **add contacts** launches the contact dialogue. The following displays the prompts received and the target details that I entered. Note that all fields are optional.

```
first_name (TEXT): Bart
middle_name (TEXT):
last_name (TEXT): Lorang
email (TEXT): lorangb@gmail.com
title (TEXT):
region (TEXT):
country (TEXT):
```

The following commands display the contacts; load the Have I Been Pwned (HIBP) credential breach module; and execute the script. The response received is identical to what you would see on the HIBP website (Chapter Eight), and is displayed after the commands below. The power with this method is that you could load hundreds or thousands of email addresses into Recon-ng and execute a search on all of them simultaneously.

```
show contacts
use recon/contacts-credentials/hibp_breach
run

lorangb@gmail.com Seen in the Bitly breach that occurred on 2014-05-08.
lorangb@gmail.com Seen in the Dropbox breach that occurred on 2012-07-01.
lorangb@gmail.com Seen in the LinkedIn breach that occurred on 2012-05-05.
lorangb@gmail.com Seen in the MySpace breach that occurred on 2008-07-01.
lorangb@gmail.com Seen in the tumblr breach that occurred on 2013-02-28.
```

We can take this type of search to another level. In Chapter Thirteen, I explained how sites such as Pastebin are often used to store user credentials after they are stolen from online services during illegal breaches. Recon-ng possesses a Paste module which will locate any pastes which contain the target email addresses, and also download the entire paste document to a text file. The following commands load the module and then execute it.

```
use recon/contacts-credentials/hibp_paste
run
```

In our example, there were no pastes that included the single email address that we have stored in our contacts. We could add more addresses through the method explained earlier, but that may be overkill for many scenarios. There is an easier way to define an email address as our target and immediately execute a search. The following commands set the source of our search as the generic email address of bob12@gmail.com, and re-execute the script. Setting this source tells Recon-ng to ignore the contacts in our database, and only focus on this single address. This type of specification of a single source works well across several modules of the application.

```
set source bob12@gmail.com
run
```

The results display the paste files that include this email within them. The raw text files for each identified paste is saved in the ".recon-ng" folder within the Home folder on the Buscador desktop. In one of these files, we can see the type of data exposed in these breaches. The following text was copied from one of the raw text files obtained. The content was slightly modified to protect the privacy of these users. The first column depicts the user number within the web service that was compromised; the second column displays the user name; the third column displays an encrypted representation of the users' passwords; and the final column confirms the email address.

```
61 MrEthic 2f6d7f5d4dbf09828c2d5427e80419b0bab69610 K@gmail.com
63 Hackfufu e5842d172949975161d8691910849ff597daa8e3 Bob12@gmail.com
64 mods88 6ac6cbceecbaaa58565805c7fa68b7a025e881d1 shaquillie668@gmail.com
```

In order to remove this temporary source (bob12@gmail.com), we can simply move to any other module. However, that email address will still be set at the "hibp_paste" module until it is overwritten with a new source. Before progressing, you should note that Recon-ng has automatically added any new contacts to our database from the previous queries. When we were searching email addresses for credentials, we had positive hits. These confirmed email addresses were added to our contacts, and are ready for additional queries. The following commands display these new contacts; load the module to identify profiles from contacts; execute the process; display the results; load the profiler module which attempts to identify additional profile names from the recently discovered data; and execute the module.

```
show contacts
use recon/contacts-profiles/fullcontact
run
show profiles
use recon/profiles-profiles/profiler
run
```

Upon completion, we now have a refined list of potential social networks of our target. As a reminder, the real power of this tool is the ability to replicate these actions across multiple targets at once. The following is a portion of the Full Contact output followed by the Profiler output.

```
lorangb              | About.me      | https://about.me/lorangb
bartlorang           | AngelList     | https://angel.co/bartlorang
lorangb@gmail.com    | Facebook      | https://www.facebook.com/bart.lorang
39267654@n00         | Flickr        | www.flickr.com/people/39267654@N00
bartlorang           | Foursquare    | https://foursquare.com/bartlorang

lorangb | WordPress | https://profiles.wordpress.org/lorangb | blog
lorangb | YouTube | https://www.youtube.com/user/lorangb/videos | video
lorangb | Tripit | https://tripit.com/people/lorangb#/profile/basic-info | travel
bartlorang | Medium | https://medium.com/@bartlorang/latest | news
bartlorang | ProductHunt | https://www.producthunt.com/@bartlorang | tech
```

We can now create our report in the same way as the first example. The following commands load the reporting module and set the customer and creator. The report will be located in the ".recon-ng" folder within the Home folder on your Buscador desktop. It will appear very similar to the report displayed in Figure 22.01.

```
use reporting/html
set CUSTOMER OSINT
set CREATOR OSINT
```

IntelTechniques

www.recon-ng.com

Recon-ng Reconnaissance Report

[-] Summary

table	count
domains	116
companies	0
netblocks	0
locations	0
vulnerabilities	0
ports	0
hosts	92
contacts	41
credentials	0
leaks	0
pushpins	0
profiles	68
repositories	0

[+] Domains

[+] Hosts

[+] Contacts

[-] Profiles

username	resource	url	category	notes	module
104592069583999251081	GooglePlus	https://plus.google.com/104592069583999251081	social		fullcontact
107828765414608142723	GooglePlus	https://plus.google.com/107828765414608142723	social		fullcontact
161353980838380954 92	GooglePlus	https://plus.google.com/109982588109119324239	social		fullcontact
86181133@N00	Flickr	https://www.flickr.com/people/86181133@N00	social		fullcontact
CallMeR0u53	Twitter	https://twitter.com/CallMeR0u53	social	Amanda :)	twitter_mentions
ErrataRob	Twitter	https://twitter.com/ErrataRob	social	Robert Graham, HODL HODL	twitter_mentions
HackingDave	Twitter	https://twitter.com/HackingDave	social	Dave Kennedy (ReL1K)	twitter_mentions
Infosecjen	Twitter	https://twitter.com/Infosecjen	social	Guv'nah	twitter_mentions
Jose.Pagliery@cnn.com	Facebook	https://www.facebook.com/jose.pagliery	social		fullcontact
Jose_Pagliery	Klout	http://klout.com/Jose_Pagliery	social		fullcontact
Jose_Pagliery	Twitter	https://twitter.com/Jose_Pagliery	social		fullcontact
KingofBigWheels	Twitter	https://twitter.com/KingofBigWheels	social	Walt	twitter_mentions
NetwarSystem	Twitter	https://twitter.com/NetwarSystem	social	Netwar System	twitter_mentions
RachelTobac	Twitter	https://twitter.com/RachelTobac	social	Rachel Tobac	twitter_mentions
SocEngineerInc	Twitter	https://twitter.com/SocEngineerInc	social	Social-Engineer, Inc	twitter_mentions
SquirrelsNaBrrl	Twitter	https://twitter.com/SquirrelsNaBrrl	social	SquirrelsinaBarrel	twitter_mentions
SultryAsian	Twitter	https://twitter.com/SultryAsian	social	SultryAsian	twitter_mentions
TeenyTinyTubers	Twitter	https://twitter.com/TeenyTinyTubers	social	Amaya Hadnagy	twitter_mentions
TheObsessedOFC	Twitter	https://twitter.com/TheObsessedOFC	social	The Obsessed	twitter_mentions
WebBreacher	Twitter	https://twitter.com/WebBreacher	social	Micah	twitter_mentions
_sn0ww	Twitter	https://twitter.com/_sn0ww	social	Snow	twitter_mentions
catmurd0ck	Twitter	https://twitter.com/catmurd0ck	social	Cat Murdock	twitter_mentions
cerealcommas	LinkedIn	https://www.linkedin.com/in/cerealcommas	social		fullcontact
chris@social-engineer.com	Facebook	https://www.facebook.com/chris.hadnagy	social		fullcontact
chrishadnagy	Gravatar	https://gravatar.com/chrishadnagy	social		fullcontact
chrishadnagy	Klout	https://klout.com/chrishadnagy	social		profiler
chrishadnagy	VideoLike	http://videolike.org/video/chrishadnagy	video		profiler
christopherhadnagy	LinkedIn	https://www.linkedin.com/in/christopherhadnagy	social		fullcontact
clutchofficial	Twitter	https://twitter.com/clutchofficial	social	Clutch	twitter_mentions

Figure 22.01: A partial Recon-ng report.



Figure 22.02: A partial Recon-ng location report.

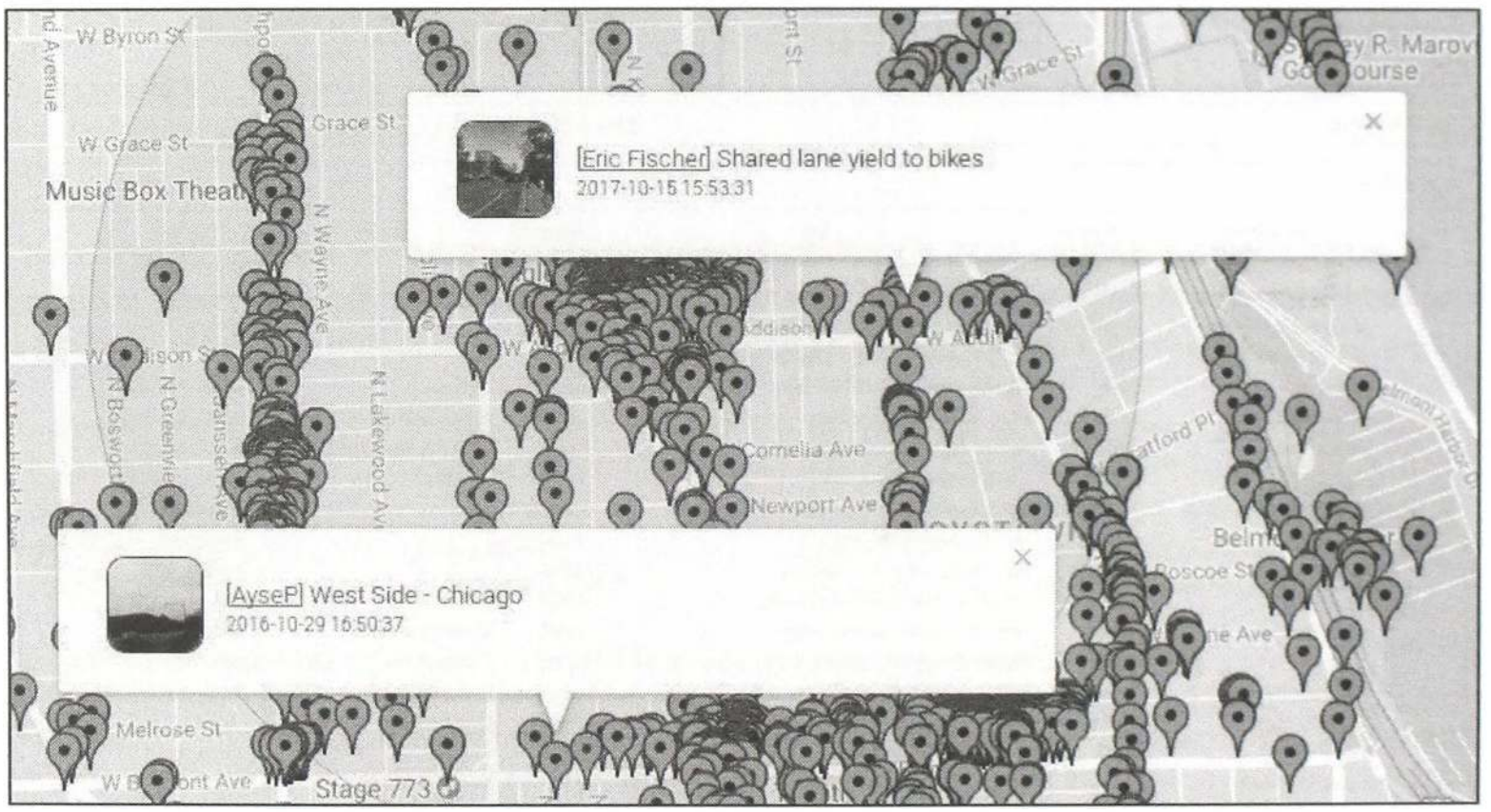


Figure 22.03: A partial Recon-ng location report.

This chapter explains only a small portion of the capabilities of Recon-ng. Please consider revisiting the modules listed at the beginning and experiment with execution of each. Overall, it would be very difficult to break the application, and any errors received are harmless. You will receive best results by requesting free API keys from the services available within the modules. If you believe a module is broken, you may just need to update your version of Recon-ng. This can be done using the instructions at the Recon-ng website, or waiting for the next Buscador update.

CHAPTER TWENTY-THREE
RADIO FREQUENCY MONITORING

Monitoring radio frequencies is considered an open source technique. It does not require the internet and the information obtained can be quite valuable. In the past, this method of collecting information was conducted using expensive hardware such as emergency receivers and console radios. Today, a few bucks or a computer can provide all of the equipment necessary to take advantage of this free information.

Hardware

There are two paths you can take in reference to hardware requirements to monitor the airwaves. Dedicated devices, such as police scanners and desktop receivers, are easier to configure. A Software Defined Radio (SDR) is more complicated and requires a computer. However, the monitoring possibilities with an SDR are greater. Overall, I recommend that a novice begin with a handheld police scanner.

Police Scanners

These affordable devices referred to as police scanners are actually capable of monitoring much more than emergency radio traffic. Practically any modern unit is capable of receiving the frequencies discussed in this chapter. More expensive units that allow for "trunked" frequencies are beneficial for monitoring government communications, but not necessary for the bulk of the radio spectrum. My personal scanner that was used for all of the methods discussed in this chapter was a Radio Shack PRO-2055. At the time of this writing, pre-owned devices were available for under $50 on Craigslist. One benefit of this type of scanner is that it can be programmed either manually or through computer software. Since the instructions for programming a scanner will vary by device, it will not be explained here. The instruction manual for the device should provide the proper instructions.

Software Defined Radio (SDR)

Some radio receivers require a computer in order to operate. These advanced systems offer many features that are not available on typical hardware scanners. These systems contain the basic hardware required to receive the frequencies and then pass data to a computer. Special software on the computer controls the hardware and allows for advanced monitoring. This can include decoding alpha-numeric pagers; receiving digital satellite imagery from weather satellites; identifying beacons from airplanes and ships; and deciphering amateur radio signals or Morse code. Explanation of these types of devices exceeds the scope of this book. If you have an interest in this, I recommend researching the topic on the internet. Beginner devices such as the

RTL2832U-based USB receiver can be purchased for $30. Higher end devices can sell for several thousand dollars. I only recommend these for tech savvy individuals.

Antennas

There are several books dedicated to theories and suggestions on proper antenna creation, alignment, and placement. The techniques discussed here do not require that level of sophistication. The supplied antenna of a police scanner should work fine. Obviously, better antennas will allow you to receive signals from a farther distance. If you find the techniques explained here successful to your investigations, then I would consider further research into superior antennas.

Frequencies

Most commercial receivers will monitor radio frequencies between 25 MHz and 1300 MHz. There are usually gaps in this coverage which will not affect the techniques described here. The techniques explained in this chapter will monitor public frequencies available on all models. All frequencies listed in this chapter are displayed in megahertz (MHz).

Family Radio Service (FRS)

If you go to the local department store and purchase a pair of two-way family radios, you will be transmitting on a public radio frequency that can be monitored with a scanner. The Family Radio Service (FRS) is a radio system authorized for use without a license since 1996. There are 14 FRS channels available which operate on 14 specific frequencies. The following table identifies these frequencies.

Channel 01 - 462.5625	Channel 06 - 462.6875	Channel 11 - 467.6375
Channel 02 - 462.5875	Channel 07 - 462.7125	Channel 12 - 467.6625
Channel 03 - 462.6125	Channel 08 - 467.5625	Channel 13 - 467.6875
Channel 04 - 462.6375	Channel 09 - 467.5875	Channel 14 - 467.7125
Channel 05 - 462.6625	Channel 10 - 467.6125	

The most common use for these radio frequencies is by families on vacation or at large events. They allow parents to keep in contact with their children. Criminals have found uses for them as well. Subjects often referred to as "spotters" use them to notify drug dealers when police are approaching a specific area. Additionally, illegal business operations such as gambling rooms and prostitution houses will use them to communicate cheaply and "anonymously". While this may afford the users some privacy protection against personal identification, the transmissions are completely public. Since the transmissions can travel several miles, the audio can be intercepted safely and without detection. Programming and monitoring these frequencies in known criminal areas may provide raw intelligence about your investigation. These are not the only frequencies to consider.

Multi-Use Radio Service (MURS)

The Multi-Use Radio Service (MURS) is an unlicensed two-way radio service that was established in 2000. The radios are capable of a range of ten miles when using decent antennas. The following table identifies the MURS frequencies.

151.820	151.940	154.600
151.880	154.570	

General Mobile Radio Service (GMRS)

Additional public frequencies, known as General Mobile Radio Service (GMRS) frequencies require a license to legally transmit audio. Most people ignore this requirement and it is seldom enforced. The radios that transmit on these frequencies can use up to 50 watts, allowing the signal to travel farther. All MURS and GMRS frequencies should be programmed and monitored in the same way as FRS frequencies. The following table identifies the GMRS frequencies.

Channel 01 - 462.550	Channel 07 - 462.700	Channel 13 - 467.650
Channel 02 - 462.575	Channel 08 - 462.725	Channel 14 - 467.675
Channel 03 - 462.600	Channel 09 - 467.550	Channel 15 - 467.700
Channel 04 - 462.625	Channel 10 - 467.575	Channel 16 - 467.725
Channel 05 - 462.650	Channel 11 - 467.600	
Channel 06 - 462.675	Channel 12 - 467.625	

Monitoring FRS, GMRS, and MURS frequencies can be crucial to several types of investigations. The following are a few scenarios that can take advantage of this method.

- Police officers can monitor criminals that use two-way radios as part of criminal activity such as drug sales in urban areas.
- Security staff can monitor families that may need emergency assistance.
- Agents can monitor groups that may be planning violence at protests.
- Investigators can monitor businesses under investigation.

Citizen Band (CB)

Many people associate Citizen Band (CB) radios with truck drivers. This is often appropriate, but truckers are not the only people that transmit on such frequencies. Since CB is low power, the receiver must be within a few miles of the transmitter. There are 40 channels available in this band. Communication on these channels may include traffic issues, witnesses to major accidents,

reports of reckless drivers, and the occasional sermon. Many state patrol vehicles include a CB radio for receiving and transmitting. The following table identifies the frequencies and channel.

Channel 01 - 26.965
Channel 02 - 26.975
Channel 03 - 26.985
Channel 04 - 27.005
Channel 05 - 27.015
Channel 06 - 27.025
Channel 07 - 27.035
Channel 08 - 27.055
Channel 09 - 27.655
Channel 10 - 27.755
Channel 11 - 27.085
Channel 12 - 27.105
Channel 13 - 27.115
Channel 14 - 27.125
Channel 15 - 27.135
Channel 16 - 27.155
Channel 17 - 27.165
Channel 18 - 27.175
Channel 19 - 27.185
Channel 20 - 27.205
Channel 21 - 27.215
Channel 22 - 27.225
Channel 23 - 27.255
Channel 24 - 27.235
Channel 25 - 27.245
Channel 26 - 27.265
Channel 27 - 27.275
Channel 28 - 27.285
Channel 29 - 27.295
Channel 30 - 27.305
Channel 31 - 27.315
Channel 32 - 27.325
Channel 33 - 27.335
Channel 34 - 27.345
Channel 35 - 27.355
Channel 36 - 27.365
Channel 37 - 27.375
Channel 38 - 27.385
Channel 39 - 27.395
Channel 40 - 27.405

Real World Application: Two truck drivers in Arkansas were engaged in a heated argument with each other over their CB radios. The argument turned into a physical altercation that resulted in severe injuries. Responding police were monitoring a CB frequency while two other truckers were following the suspect. This communication helped police identify and arrest the suspect.

Marine Channels

Frequencies for marine use are not titled in the same fashion as other groups. Some channels are skipped and others are amended with the letter "A" at the end. There is a general understanding within marine circles of the type of radio traffic on each channel, but there is no enforcement of these designations. For example, many channels are reserved for commercial traffic only, but some disregard this rule. Many people disregard scanning this type of radio traffic because they do not live near waterways. Some criminal groups have purchased portable "walkie-talkie" style marine radios in non-water areas and communicate without fear of authentic marine vessels hearing their traffic. I recommend programming all of the frequencies into your scanner in the same bank as other civilian frequencies. The following page contains a table of channels and frequencies used for marine communications.

Real World Application: While assisting a law enforcement agency during a large protest, I monitored marine frequencies for suspicious traffic. Because we were near a large body of water and several marinas, I assumed there might be intelligence to be gleaned. I immediately identified a small group of people communicating with portable marine radios. They were using them to coordinate meeting locations for large groups that were arriving by land and sea. While this intercepted traffic did not pose a threat, the intelligence was beneficial to the mission to keep everyone safe.

01A	156.050	20A	157.000	72	156.625
05A	156.250	21A	157.050	73	156.675
06	156.300	22A	157.100	74	156.725
07A	156.350	23A	157.150	77	156.875
08	156.400	24	161.800	78A	156.925
09	156.450	25	161.850	79A	156.975
10	156.500	26	161.900	80A	157.025
11	156.550	27	161.950	81A	157.075
12	156.600	28	162.000	82A	157.175
13	156.650	63A	156.175	84	161.825
14	156.700	65A	156.275	85	161.875
15	156.750	66A	156.325	86	161.925
16	156.800	67	156.375	87	157.375
17	156.850	68	156.425	88A	157.425
18A	156.900	69	156.475	AIS 1	161.975
19A	156.950	70	156.525	AIS 2	162.025
20	161.600	71	156.575		

Room Monitors

There are several ranges of frequencies assigned by the Federal Communications Commission (FCC) that are used by personal devices. These include a group of frequencies designated for one-way monitoring devices often used as "baby monitors". These devices come in pairs. One unit is a transmitter that is placed in a room to be monitored, such as a nursery. The second unit is a receiver that can be up to 1000 feet away that will broadcast the audio from the other unit. Often, these include a switch on each unit to switch between two or three different channels. This is because all of these devices use the same frequencies. If you received interference from a neighbor's device, you could switch the channel. The channels assigned to these units fall in three ranges. The 49 MHz models are older and cheaper units that operate on a frequency similar to older cordless phones. The following frequencies are assigned to these units.

49.300	49.830	49.845	49.860	49.875	49.890

The 900 MHz models are more popular and are also used in wireless video cameras with audio. Houses that have consumer grade wireless surveillance cameras operate within this band. Often these devices transmit audio and video to a base station that is connected to a television in the house. I have seen many of these during search warrant executions. I believe any good tactical operations plan should include a sweep of this frequency range before execution of a search warrant. This can identify the presence of wireless systems used to provide early notice of law enforcement at the door. This can be a huge officer safety concern. There are too many frequencies available to these devices to list here. Instead, I recommend a scan of the following two ranges. Be warned that you may encounter cordless telephones within this range. If you do, block the frequency in order to avoid any law violations.

902.000 MHz - 908.000 MHz 923.000 MHz - 928.000 MHz

The 2.4 GHz models are often encrypted and change frequencies sporadically. Additionally, most basic scanners cannot monitor this range. Only expensive highly specialized devices can accurately monitor this traffic.

Real World Application: A police department in a Chicago suburb executed a search warrant on the home of a child predator. The suspect heard the entry from a basement room where he possessed the receiver of a wireless audio room monitor. This early warning allowed him time to destroy evidence and flee through a basement window. A standard frequency scanner could have notified the officers that a wireless system was present. This may have prompted a closer analysis for cameras and modifications could have been made to the operations plan.

Wireless Microphones

Wireless microphones or "cordless" microphones can be found transmitting throughout the radio frequency spectrum. They are most often in the 42 MHz, 70-74 MHz, 170-220 MHz, and 580-800 MHz ranges. Many modern professional wireless microphone systems are frequency agile and tunable to different frequencies. Wireless microphone power levels are normally very low in order to reduce the potential for interference. Clear reception requires close proximity or use of directional antennas. The following frequencies should receive audio from most of the existing wireless microphone systems. If you expect to receive a wireless transmission, and none of these frequencies are active, you should search by "Nearby Frequencies".

169.445	73.3000	178.2000	187.6000	199.6000
169.505	73.6000	178.6000	188.4000	200.4000
170.245	75.1000	179.2000	190.2000	202.2000
170.305	75.5000	180.8000	192.8000	202.6500
171.045	75.7000	181.2500	193.2000	204.8000
171.105	75.9000	181.6000	193.6000	205.6000
171.845	82.5000	182.4000	194.4000	206.3500
171.905	82.8000	183.2000	195.2000	206.4000
72.1000	83.8000	183.4000	195.4000	208.2000
72.3000	86.8000	184.2000	195.8000	208.6500
72.5000	174.8000	184.6000	196.2000	210.8000
72.7000	175.2500	184.8000	196.6000	211.6000
72.9000	175.6000	185.2000	197.4000	212.4000
73.1000	176.4000	186.8000	198.7500	

Real World Application: While attending a computer security conference, a non-scheduled presentation was offered to select attendees in reference to new exploits being used on corporate networks. Since I did not receive an invitation, I was not allowed entry. I connected my headphones to a portable scanner and began scanning the common ranges used by wireless microphones while sitting in the hallway outside of the closed-door session. Within moments, I had discovered the frequency used by the speaker's wireless lapel microphone and was able to listen to the entire presentation. It is likely that I was able to hear the speaker clearer than most of the live audience since I had the isolated microphone feed.

Hotels and Convention Centers

Most hotels use some type of two-way radio systems as part of their daily operations. Many larger hotels in major cities use a group of channels and isolate traffic for maintenance, housekeeping, security, and valet. This traffic can identify valuable information during a targeted investigation. A security protection detail that is assigned to a high-profile subject may want to monitor the frequencies of the hotel where the subject is staying. The traffic may identify employees discussing the subject and announce vulnerabilities including the assigned room number and specific requests by the subject. The vehicle information may be broadcasted on the valet channel and gossip may be heard on the operations channels. This information could be devastating if used maliciously. Additionally, traffic on the hotel security channel could identify a threat to a target before the security detail is involved. Law enforcement analysts should begin monitoring a specific hotel's frequencies the moment a serious event occurs at the location such as an explosion, hostage situation, bomb threat, homicide, armed robbery, etc. This live information can be vital to the investigation. Hotels do not use a standard set of frequencies or band on the radio spectrum. Discovering the frequencies to monitor is easy thanks to Radio Reference.

Radio Reference (radioreference.com)

This is the most complete collection of current frequencies assigned to government agencies and private businesses on the internet. It was discussed briefly in Chapter Seven and online monitoring of emergency communications was explained. You can also use this database to identify practically any active frequency. Basic search methods can be completed without a premium account. The search field in the upper right portion of every page can locate a set of frequencies based on a business name. Additionally, you can browse by location and identify all businesses in a specific area. This section will explain both methods.

A search of "Chicago Palmer House" on Radio Reference identifies all of the frequencies used by the Palmer House Hilton hotel in downtown Chicago. This hotel is often used by dignitaries and public officials. Figure 23.01 displays the results including the assigned call sign, frequency, and number of units allowed to access the frequency. Within these ten frequencies are the channels assigned to security, housekeeping, maintenance, and administration. The valet at this

location now uses Nextel cellular service. These active channels can be monitored from several miles away using a basic scanning device.

You can also browse the Radio Reference database to locate frequencies of interest. Clicking the "Databases" link at the top of every page will present an interactive United States map. You can click any state and will be presented all of the counties in that state. Selecting a county will present the options of government and business frequencies for that area. Figure 23.02 displays the results for Guaranteed Rate field. It identifies thirteen frequencies assigned to Security, Parking, Maintenance, Operations, and others. This could be used to quickly begin monitoring these frequencies if a threat or catastrophe occurred at that location.

Entity	Callsign	Frequency	Units
THOR PALMER HOUSE HOTEL DBA/PALMER	WQEB345	451.28750	50
THOR PALMER HOUSE HOTEL DBA/PALMER	WQEB345	451.58750	50
THOR PALMER HOUSE HOTEL DBA/PALMER	WQEB345	456.28750	50
THOR PALMER HOUSE HOTEL DBA/PALMER	WQEB345	456.58750	50
THOR PALMER HOUSE HOTEL DBA/PALMER	WQEB345	461.68750	50
THOR PALMER HOUSE HOTEL DBA/PALMER	WQEB345	461.96250	50
THOR PALMER HOUSE HOTEL DBA/PALMER	WQEB345	462.21250	50
THOR PALMER HOUSE HOTEL DBA/PALMER	WQEB345	466.68750	50
THOR PALMER HOUSE HOTEL DBA/PALMER	WQEB345	466.96250	50
THOR PALMER HOUSE HOTEL DBA/PALMER	WQEB345	467.21250	50

Figure 23.01: A Radio Reference search result.

Frequency	License	Type	Tone	Alpha Tag	Description
461.45000	WQAU450	RM	67.0 PL	CWS Security	Security - Main (as of June 2010)
462.05000	WPXR683	RM	732 DPL	CWS Ops F-3	Guest Relations Operations [F-3]
461.20000	WPLI617	RM	67.0 PL	CWS Parking	Parking [F-6]
456.56250		M	051 DPL	CWS 456.5625	Food-Beverage service
463.72500	WPLL482	RM	466 DPL	CWS Maintnce	Maintenance
464.28750		M	67.0 PL	CWS Ticketng	Ticketing
464.51250		M	226 DPL	CWS Food	Food
464.55000		M	047 DPL	CWS Ops46455	Operations
464.67500	WQDD864	RM	223 DPL	CWS Concessn	Sportservice - Concessions
464.81250		M	466 DPL	CWS Food	Food
464.83750		M	051 DPL	CWS Janitor	Janitorial
464.95000	WPLL482	RM	67.0 PL	CWS Ops D	Operations (infrequently used)
464.75000		RM	67.0 PL	CWS Sec old	Security - Main (old)

Figure 23.02: A Radio Reference search result.

Retail Businesses

The next time you are shopping at a large store such as Walmart or a clothing store such as Old Navy, pay attention to the employees. Most of them will either have portable radios in their pockets or wireless headsets. This is how they communicate to request more cashiers, announce a lunch break, verify a price, and direct employees to different areas. When you order food at the drive-through at a fast food restaurant, you are probably talking through a two-way radio device.

When you are at a large concert, there are several people around you communicating through radio systems. This will include servers carrying drink orders, security, and even the backstage crew working with the musicians. This is all available to you through radio frequency monitoring. Practically anywhere your investigation takes you will present possibilities in intelligence collection through these methods. The frequencies of the businesses in your location should be available on Radio Reference.

Real World Application: Detectives in a St. Louis suburb were investigating a report that an employee at a local fast food restaurant was distributing cocaine while working. Covert Officers inside the business never noticed any unusual activity. After monitoring the drive-through frequency, they overheard several orders for a specific non-food item that was not on the menu. They later determined that this was the code word for one gram of cocaine. Several arrests were made.

News Media

During large investigations, various print and video news outlets will occasionally obtain information before the police do. This may be through witness reports or diligent efforts of a skilled reporter. Usually, it is due to the financial resources of the media company. These companies are quick to send a helicopter to the scene of a crime or a news van to a victim's house to capture video footage. While most ground reporters use cellular telephones to communicate with the news room, helicopters use radio frequencies. These powerful transmissions can be received several miles away. The frequencies for your area can be found on Radio Reference.

Real World Application: An Illinois police department was investigating a missing person that was last seen near a rock quarry. A news helicopter was flying overhead filming officers walking in the bottom of the quarry. The pilot observed what he thought may be the victim's body and began transmitting this through the radio system to the news desk. An alert officer monitoring the helicopter from a portable radio heard the report and determined the location of the body based on the pilot's description to his co-workers.

Emergency Communications

One last obvious group of interesting frequencies is that of emergency personnel such as police, fire, and EMS. Law enforcement officers should keep the frequencies of surrounding agencies stored in their scanners. This can provide valuable information such as pursuits, major incidents, and crimes related to their own jurisdiction. Those not in law enforcement can also use this technique for intelligence gathering. News reporters would want to monitor all local emergency channels for immediate notification of the next hot scoop.

Nearby Frequencies

If you cannot locate the specific target frequency to load into your scanner, you have one other option. Most modern devices include a feature that will quickly scan all of the major bands of frequencies for any that are in your immediate area. Uniden radios refer to this as the "Close Call" option while others label it as "Signal Stalker". Basically, it scans for the strongest frequencies and ignores weak signals. This may be the only way that you can identify the frequency that you are seeking. This will often identify local frequencies of interest that are not listed on Radio Reference or are unlicensed.

Wireless Video Cameras

Surveillance systems with a Digital Video Recorder (DVR) are very common in homes and businesses. An abundance of electronics from China has made these affordable for anyone. In the past, the solution was to install cameras that were wired directly into a DVR. This caused a web of hidden wires inside of walls and ceilings. Today, most people choose a wireless system that broadcasts live video on the 900 MHz, 1.2 GHz, or 2.4 GHz spectrum. The consumer models rarely encrypt the signal, and any generic video receiver can view the live stream. The FCC has assigned specific blocks of frequencies for the wireless channels used by these devices. Because of this, new handheld receivers have been created that can scan all of these frequencies and display any wireless video on a small screen. These range in price from $400 to $500. A Google search for "wireless video scanner" will present many options. A decent unit with a quality antenna will display any wireless video being transmitted within an entire residential block. These can also be used to detect unauthorized hidden video devices.

Always use discretion when scanning for wireless video signals. While completely legal to view, using the collected video to harass, intimidate, stalk, burglarize, or defraud someone is obviously illegal and unacceptable. Many people have been arrested for abusing this technology.

Broadcastify (broadcastify.com)

Broadcastify is an online portal for scanner enthusiasts. It contains the most comprehensive database of radio frequencies that can be monitored on personal radios and scanners. The forums are very active with conversations about emergency radio traffic, hardware reviews, amateur radio, and frequency monitoring software. This can all be interesting content to search, but will rarely provide valuable intelligence. The real interest for OSINT analysts is the live and archived audio. The live audio feeds are free and do not require an account to access them. You will be presented a map to which you can navigate to the state and county of interest. This will present all of the live audio feeds of public scanner frequencies that can be monitored over the internet. These live audio feeds are provided by individual listeners within the area of the frequency coverage. A user leaves a personal scanner on the frequency and broadcasts the audio through a computer with internet access. You will find audio feeds in all parts of the country. This includes large cities, small towns, and rural counties. This can provide immediate information about live

events anywhere. You no longer need to be in a specific location to receive radio traffic from that area. For historical content, information can be obtained from the archives.

Most live audio feeds have an option within the player window of "Feed Archives". This will take you to a menu that will provide several dates from which to choose, a time frame of a chosen date, and an embedded player that will play the radio feed from the time specified. Additionally, you can choose the "save as" feature to download a copy of the communication as an MP3 file for archiving. This will provide the emergency services radio traffic during an event of interest. The archive function requires a user account.

Web SDR (websdr.org)

The previous technique connected you to shared radio scanners set to a specific frequency. There is another community that allows you to control the target frequency. A Web SDR is a Software-Defined Radio receiver connected to the internet. It allows many listeners to listen and tune it simultaneously. SDR technology makes it possible that all listeners tune independently, and thus listen to different signals. This is in contrast to the many classical receivers that are already available via the internet. This page will direct you to over 100 online receivers waiting for your connection. The list will identify the location, frequency range, and antenna in use. The world map at the bottom of the page will help you quickly locate receivers in your targeted area. This can also be used to monitor a digital communication over an analog signal. This tone can be converted to the text that it is sending to another device. I have used this in the past to monitor marine, citizen band, and ham signals in geographical areas of an active investigation while I was states away. Figure 23.03 displays monitoring of a radio frequency through a receiver in San Francisco while I was in Washington, D.C. The "Start" button in the center allows constant recording of a frequency while monitoring.

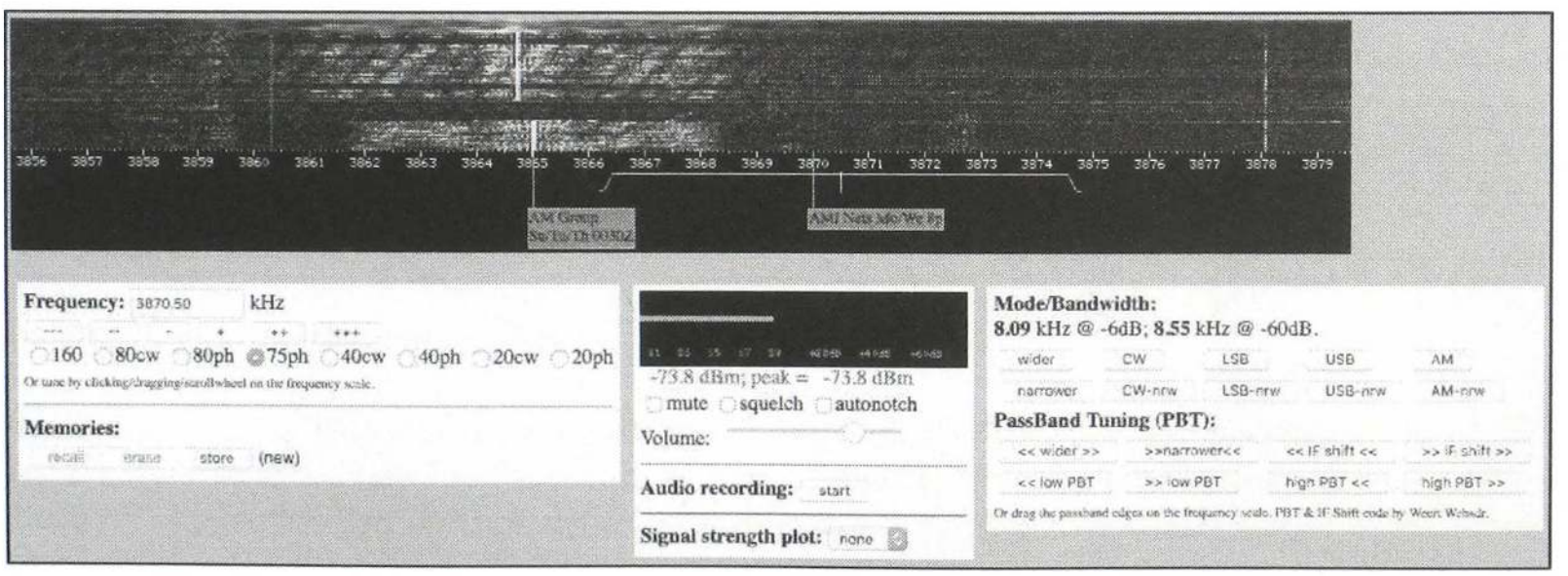


Figure 23.03: A remote frequency monitor on Web SDR.

CHAPTER TWENTY-FOUR
OSINT WORKFLOW PROCESSES

I have conducted numerous OSINT training programs over the past few years. Regardless of the audience, I receive one question at every event.

"Is there a standard process or workflow for OSINT?"

My short answer was always "no". I had always looked at each investigation as unique. The type of investigation dictated the avenues and routes that would lead me to valuable intelligence. There was no cheat-sheet that could be used for every scenario. While I still believe that there is no template-based solution for this type of work, I now admit that some standards can be developed. This chapter will display my attempt at creating workflows that can quickly assist with direction and guidance. These documents are presented in six views based on the information being searched. Each example should be considered when you are researching the chosen topic. The categories are Email, User Name, Real Name, Telephone Number, Domain Name, and Location.

Each example will try to show the standard path that I would take when provided the chosen type of data, such as an email address. The goal with my investigations is to get to the next topic. For example, if I am given an email address, my goal is to find any user names and real names. When I have a user name, my goal is to find any social networks and verify an email address. When I have a real name, the goal is to find email addresses, user names, and a telephone number. When I have a telephone number, my goal is to verify the name and identify a physical address and relatives. When I have a domain name, my goal is to locate a real name and address. The cycle continues after each new piece of information is discovered.

Each example will identify only the services used. It will not display the actual address to navigate to the website. However, every method listed within these charts is explained throughout this book. These documents do not contain every avenue that may provide good information. They only display the most beneficial resources. Think of them as the obvious steps to take when you receive a target to search. The results of your queries can lead you to more places than can display on a single page in this book. These are just the first priorities.

Consider the Email flowchart presented in two pages from now. The written translation of this would be to take the email address and search it within the Hunter email validation tool. Next, conduct searches of the address within quotation marks on the main search engines. After that, check the compromised databases and all options on the IntelTechniques Custom Email Search Tool. These options are likely to lead to a presence on social networks. The Facebook will likely identify a user name, employer, and real name. Following these to the bottom of the chart encourages you to conduct the email assumptions mentioned previously, which you can verify and start over with the newly acquired information. You would then continue through the

remainder of the chart. If you find the following information beneficial, you are welcome to download digital copies from my website at inteltechniques.com/data/workflow.zip. I also recommend visiting osintframework.com. While it is not a traditional workflow, it does provide numerous online resources within an interactive tree. Many of the websites mentioned here are replicated on osintframework.com, which was created by Justin Nordine.

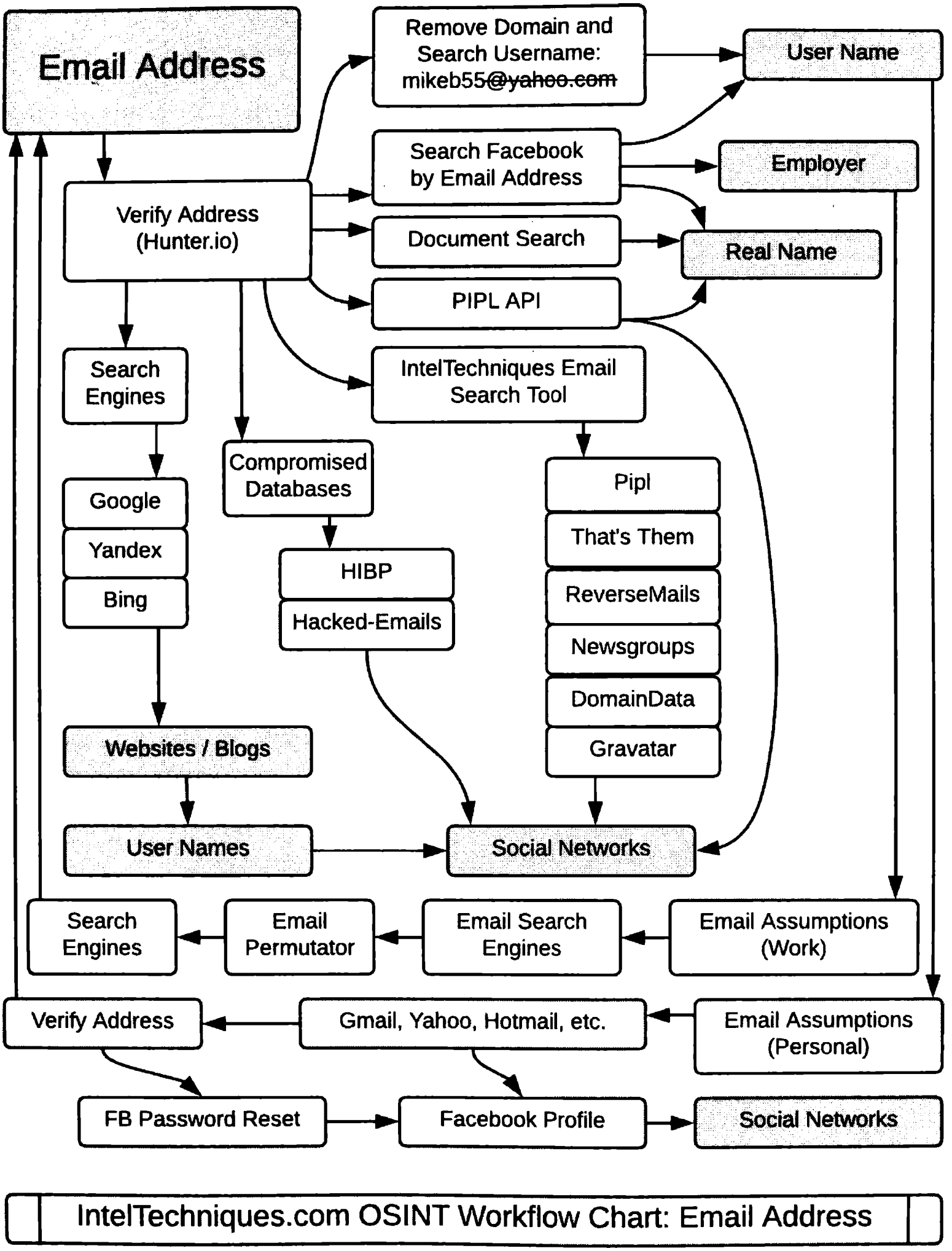
I believe that all of these will always be a work in progress. As everything else in OSINT changes, these will too. I will try to keep them updated on the website. If you have suggestions, I am honored to receive and apply them. If you would like to create better formulas, I encourage you to get creative. I used the website LucidChart.com to create each of these. I also made all of these public within the LucidChart website so that you can take advantage of my starting point. The following links will connect you to a live environment that will allow you to replicate the work in seconds. If you would like a similar service without the requirement of a registered account, please consider **MindMup** (mindmup.com).

Email Address: lucidchart.com/invitations/accept/5282ad5a-b0dc-4442-a4a5-4a440a00dd05
User Name: lucidchart.com/invitations/accept/5282ad70-58dc-4546-8758-0a460a00c875
Real Name: lucidchart.com/invitations/accept/5282ad8b-c4d0-4db3-98f2-25d00a00c875
Telephone: lucidchart.com/invitations/accept/5282ad9a-64a4-4435-9073-3ce80a00c875
Domain Name: lucidchart.com/invitations/accept/5282acc9-f324-43b2-af40-04c00a00c875
Location: lucidchart.com/invitations/accept/9d446294-580e-49ba-a88f-2437cc392b6f

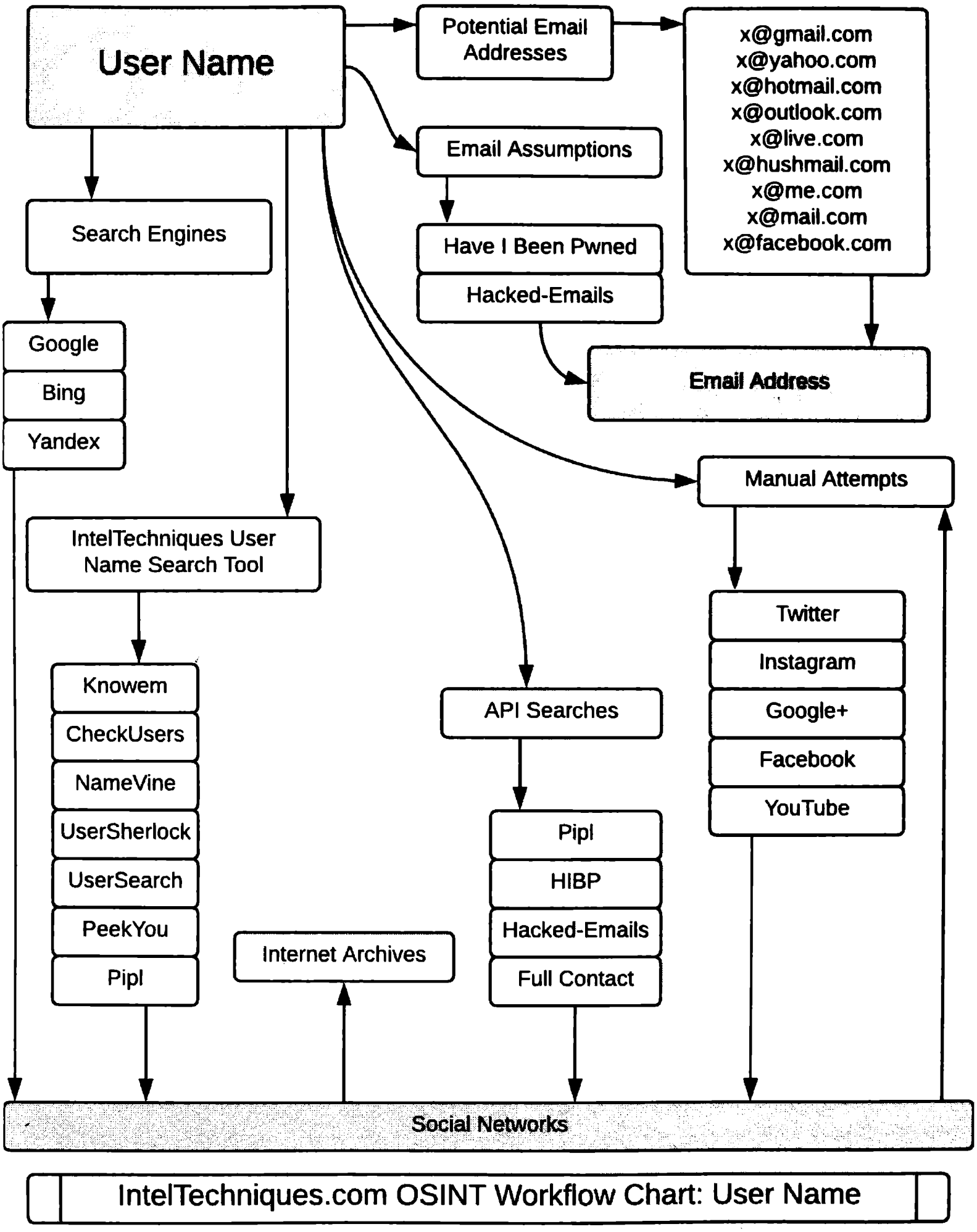
Many readers have requested practical exercises in order to test their OSINT skill. I agree that this would be helpful, but maintaining active and accurate online demonstrations with live data can be overwhelming. Instead, I encourage you to test your skills with real data, unknowing to the target. Consider the following scenarios, and use the flowcharts here as a guide.

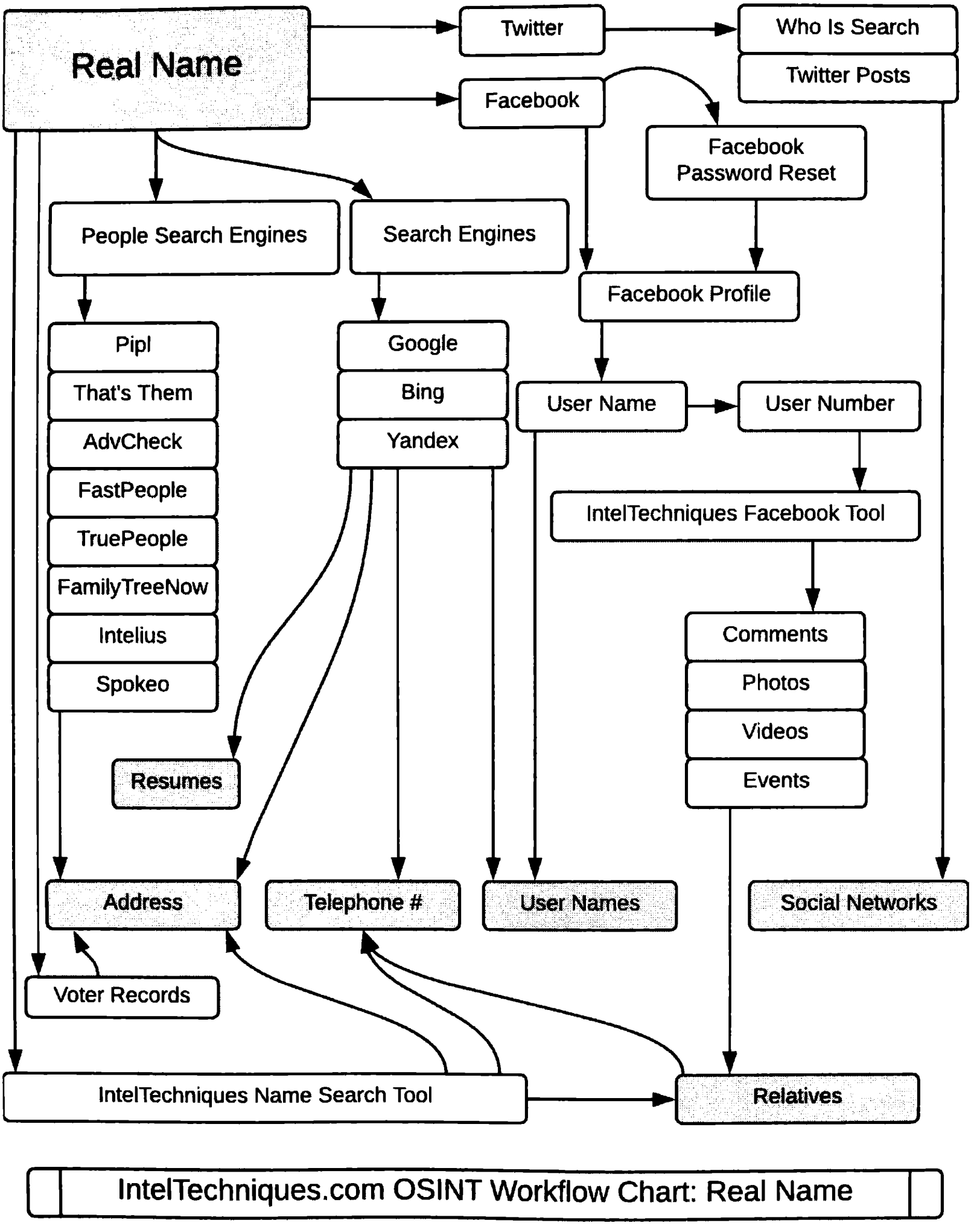
Zillow: Pick a random home and find all info about the previous owners
Wrong Number (incoming): Reverse-search it, text them their details
Wanted Criminals: Locate any significant others' online profiles with photos
Waiter/Waitress: Research your server from dinner last night and identify their vehicle
AirBnB: Locate all details about a host (owner) and email them directly
Radio: Pick a morning "Happy Birthday" target, obtain full DOB and relatives' comments online
Reviews: Find 5 people that have patronized a local business and locate their home addresses
Game Show Contestant: Identify full address, phone number, photos, and relatives
Newspaper: Choose a person quoted in today's newspaper and identify their social networks
News: When a local Facebook comment is cited, explore the hidden data about the person
Library: Locate an employee's Amazon wish list and buy them the book he or she wants (creepy)

This list could grow for many pages. Overall, there are endless targets available that provide the best practice possible for exercising these techniques. This practice will increase the confidence in your research during an actual investigation. The hard part is not disclosing what you find to them. While you may think they will be impressed with your new skills, they won't. Trust me...

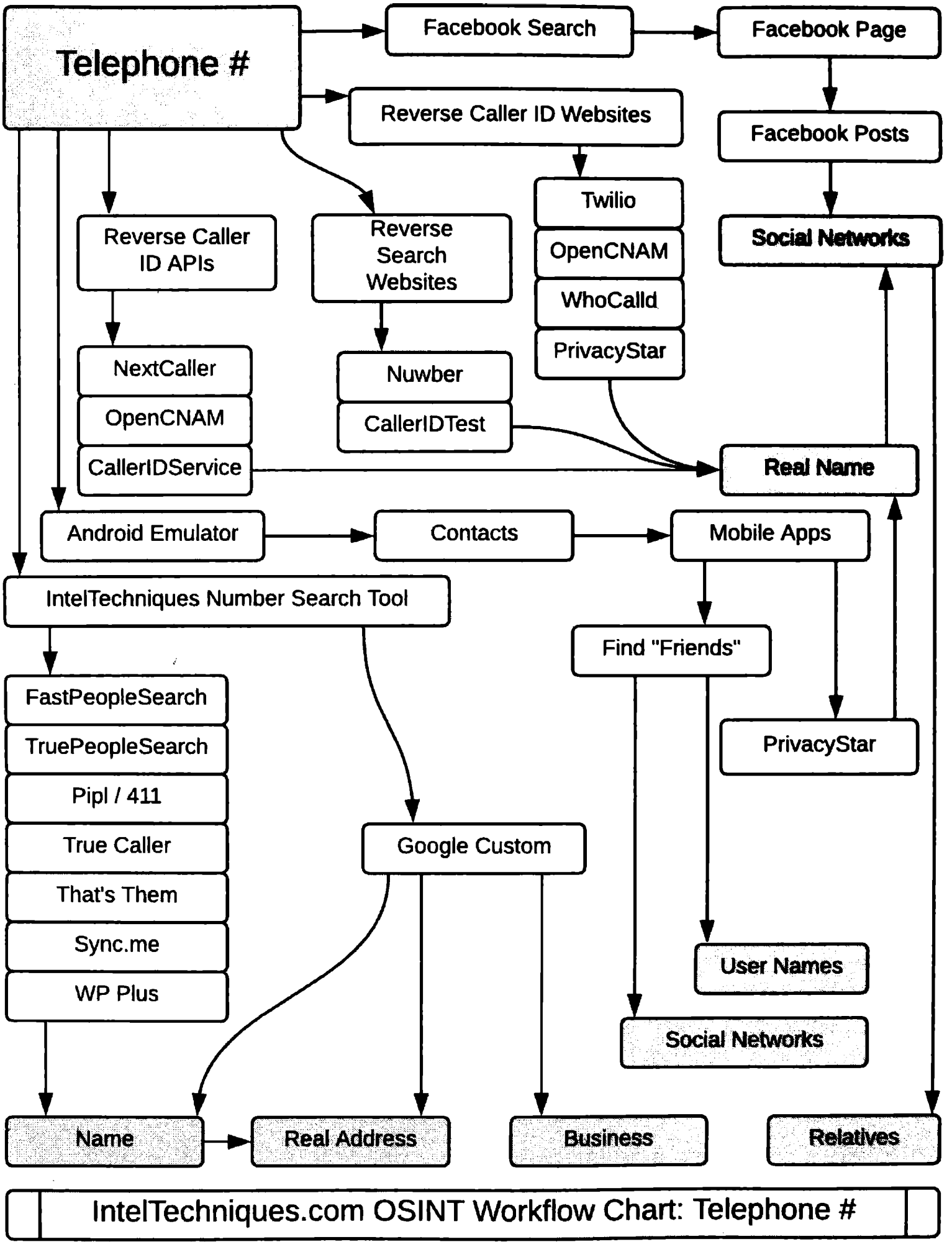


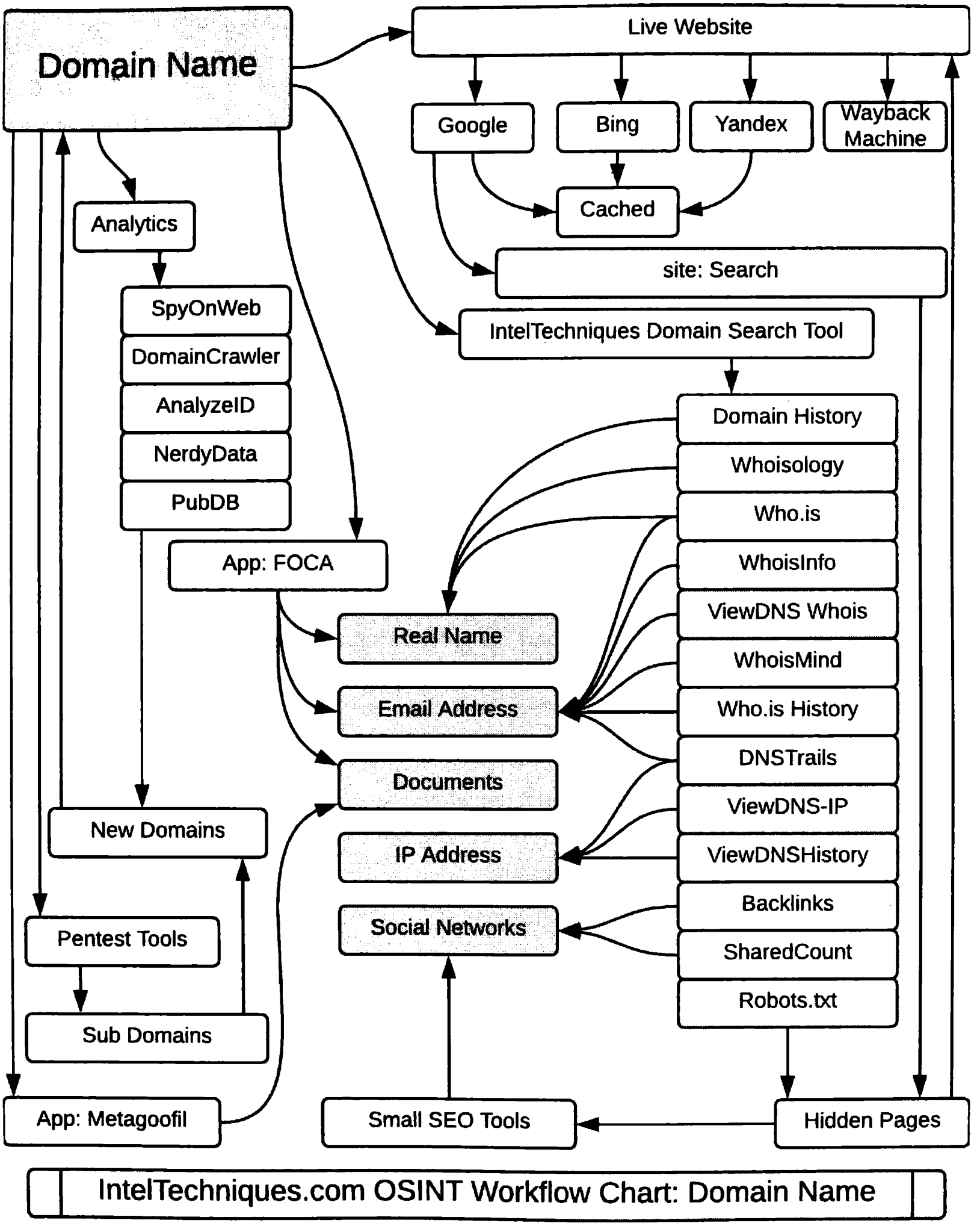
IntelTechniques.com OSINT Workflow Chart: Email Address


User Name
Potential Email Addresses
x@gmail.com
x@yahoo.com
x@hotmail.com
x@outlook.com
x@live.com
x@hushmail.com
x@me.com
x@mail.com
x@facebook.com
Email Assumptions
Search Engines
Have I Been Pwned
Hacked-Emails
Email Address
Google
Bing
Yandex
Manual Attempts
IntelTechniques User Name Search Tool
Twitter
Instagram
Google+
Facebook
YouTube
Knowem
CheckUsers
NameVine
UserSherlock
UserSearch
PeekYou
Pipl
API Searches
Pipl
HIBP
Hacked-Emails
Full Contact
Internet Archives
Social Networks
IntelTechniques.com OSINT Workflow Chart: User Name

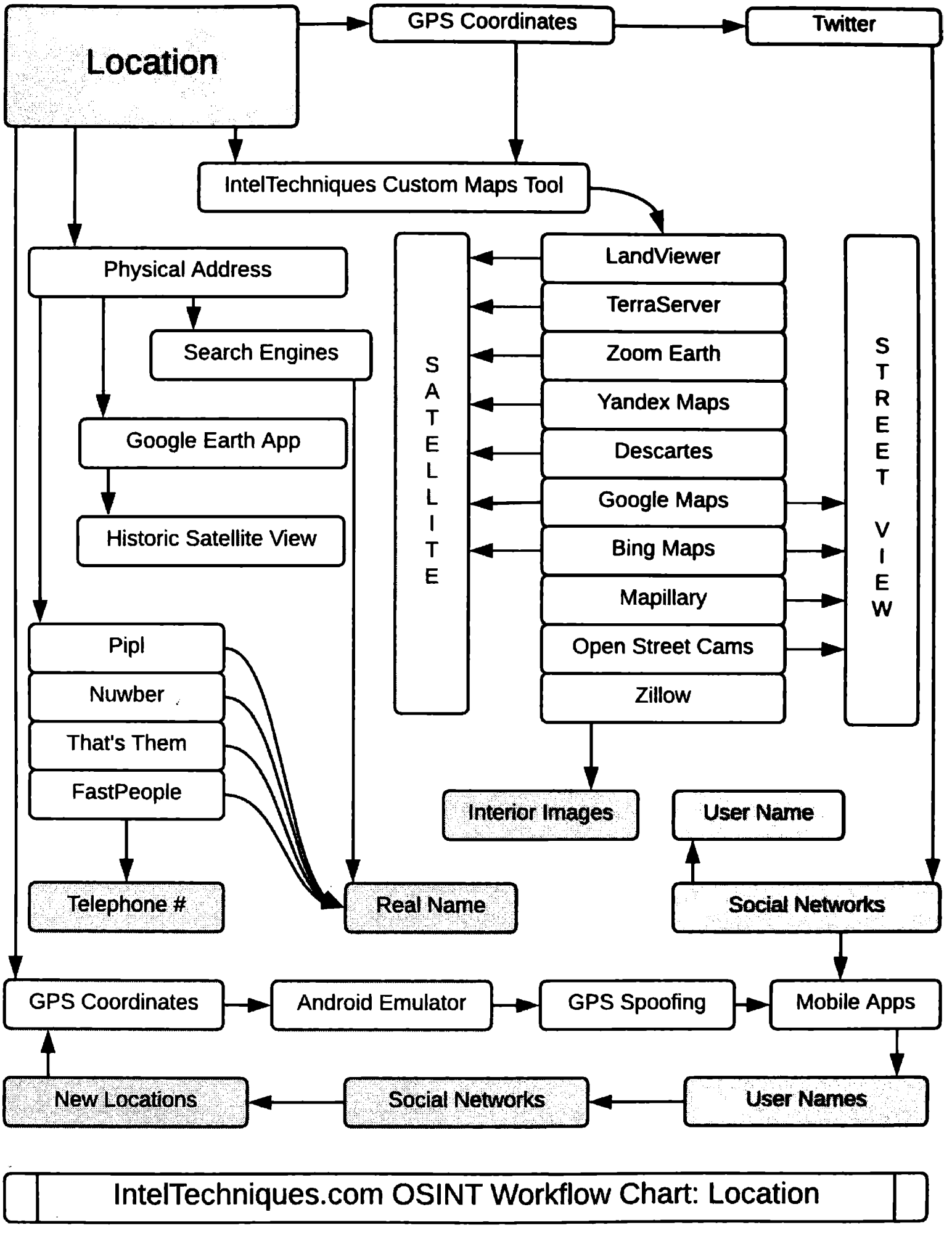


IntelTechniques.com OSINT Workflow Chart: Real Name


Telephone #
Facebook Search
Facebook Page
Facebook Posts
Social Networks
Reverse Caller ID Websites
Twilio
OpenCNAM
WhoCalld
PrivacyStar
Reverse Caller ID APIs
Reverse Search Websites
NextCaller
OpenCNAM
CallerIDService
Nuwber
CallerIDTest
Real Name
Android Emulator
Contacts
Mobile Apps
IntelTechniques Number Search Tool
Find "Friends"
PrivacyStar
FastPeopleSearch
TruePeopleSearch
Pipl / 411
True Caller
That's Them
Sync.me
WP Plus
Google Custom
User Names
Social Networks
Name
Real Address
Business
Relatives
IntelTechniques.com OSINT Workflow Chart: Telephone #



IntelTechniques.com OSINT Workflow Chart: Domain Name


Location
GPS Coordinates
Twitter
IntelTechniques Custom Maps Tool
Physical Address
Search Engines
Google Earth App
Historic Satellite View
Pipl
Nuwber
That's Them
FastPeople
Telephone #
Real Name
SATELLITE
LandViewer
TerraServer
Zoom Earth
Yandex Maps
Descartes
Google Maps
Bing Maps
Mapillary
Open Street Cams
Zillow
STREET VIEW
Interior Images
User Name
Social Networks
GPS Coordinates
Android Emulator
GPS Spoofing
Mobile Apps
New Locations
Social Networks
User Names
IntelTechniques.com OSINT Workflow Chart: Location

CHAPTER TWENTY-FIVE
DOCUMENTATION

By now, hopefully you have learned many new ways of collecting online information. Usually, possessing the information is not enough. At some point, you are likely going to be required to document your findings and provide a summary report. For law enforcement, this may be an agency report within a records management system. For the private sector, it may be a detailed memo for internal use. Private investigators may need to create their own summary and provide it to the client. There is no standard template that fits all of these scenarios. However, there are some guidelines that may help you create the most appropriate documentation of your online investigation. This chapter attempts to convert your online evidence into digestible reports.

First, we should separate into two categories: Government and private sector. These are two very different groups with unique requirements. I will start with the private sector. Since I began using OSINT techniques as part of my investigations, I have seen a drastic shift in its use within the private sector. Many corporations now possess an OSINT unit. It may have a different title, and it may consist of only one person, but you know who you are. During my live and online trainings, I meet many people that solely focus on OSINT within their job. Companies have accepted that internet-based research is required as part of the overall investigative role. If you are responsible for conducting online investigations, detailed reports are as crucial as the content found. In my early years of OSINT work, my documentation method was simple. Open Microsoft Word, add a few case descriptors, and start hunting. As I found interesting information, I would include a brief statement or two and then add a screen capture. I would then repeat until the case ran dry. I am embarrassed to look at these reports today.

I no longer document any findings until the end of my case. As long as you are using FireShot (free) or Hunchly ($), every screen capture is date and time stamped. You do not need to write a report concurrently with the preservation of your findings. I prefer to focus all effort toward locating the evidence versus explaining each process. After you have archived all of the information with your collection utility of choice, then you should start the documentation. I believe the first consideration for your OSINT report should be the "Executive Summary". This document should be limited to one page, and it should only focus on the absolutely vital evidence found during your investigation. I limit this to one page because you will lose your audience quickly if it is any longer. Think of it as the "elevator pitch". You have only a few minutes on the elevator with the person that needs to know what you have found. There is no time to explain how a reverse-video search works or about the ways you can exploit a Facebook account with a user ID number. You only have time to tell the boss what he or she needs to know. This may be a summary of the facts proving that an employee is stealing from the company; identification of the person that posted violent threats toward a client; or the most damaging evidence that enforces your opinion that a potential employee not be hired. You have one page, so take advantage of every inch. The following page displays an actual modified Executive Summary.

Investigation Number: 2017-54887
Investigator: Michael Bazzell

Date: November 17, 2017
Suspect: John Davis

On November 12, 2017, I was requested to conduct an investigation into John Davis, a former employee that has made claims against your company in reference to being injured on the job. Upon completion of my investigation, I have collected 143 relevant screen captures of online evidence that validates your resistance to his claims. The following pages include detailed analysis of the online content, but this cover page summarizes the most damaging evidence found that contradicts his previous deposition testimony. I find the following facts to be most useful.

- On June 13, 2017, Mr. Davis claimed to sustain a back injury while working at the South Plant on June 12, 2017. He was sent home on the 13, and has not returned to work since.

- On June 11, 2017, Mr. Davis received an invitation through a Facebook Event from his sister, Jane Davis to attend a retirement celebration at Lake of the Ozarks, Missouri. This event was scheduled to take place on Saturday, June 17, 2017.

- According to his deposition, Mr. Davis' chiropractor is Neil Stevens in Ladue, Missouri. Dr. Stevens ordered Mr. Davis on bedrest until further notice on June 14, 2017, and continued this order on July 27, 2017.

- On June 18, 2017, Jane Davis posted numerous photos to her Facebook profile, including an image of Mr. Davis lifting a keg of beer above his head.

- On June 27, 2017, Mr. Davis participated in a 5K race in St. Louis. Online records indicate that he placed 113 out of 1,237 with a time of 00:28:16.

- On August 12, 2017, Mr. Davis posted a review on Amazon for a steel rack for mounting onto an ATV. The review included, "This thing survived five miles of rough terrain last week, and my rifles didn't get one scuff ... I will never go hunting without this on my 4-wheeler".

- Dr. Stevens's daughter possesses an Instagram profile indicating that she is best friends with the daughter of Mr. Davis. Mr. Davis' daughter attended a sleepover at Dr. Stevens' home on March 12, 2017. Dr. Stevens is an avid hunter, as is Mr. Davis. On September 22, 2017, Mr. Davis and Dr. Stevens placed second in a duck hunting competition in Boone County, Missouri. Online records confirm they were on the same team.

The following pages represent the evidence of these findings. Please note that all screen captures were preserved in digital format, and are included within the DVD attached to this report.

The previous example provided just enough detail to give an overall synopsis of the case. Most clients will not read past this page until necessary. They may thumb through the entire report and look at any screen captures, but ultimately, they possess the information they need. Obviously, you still need to provide your evidence, which is what we will do in the next portion of the report. I use a specific "Suspect Details" template, but you should create your own that best represents your needs. Consider my example below.

Investigation Number:		Date:
Investigator:		Suspect:
Full Name:	Age:	DOB:
Home Address:		Telephone:
Mailing Address:		Telephone:
Spouse:		
Child # 1:		
Child # 2:		
Suspect Email Addresses:		
Spouse Email Addresses:		
Child # 1 Email Addresses:		
Child # 2 Email Addresses:		
Suspect User Names:		
Spouse User Names:		
Child # 1 User Names:		
Child # 2 User Names:		
Suspect Social Network Profiles:		
Facebook:		Twitter:
Instagram:		Google:
Other:		Other:
Souse Social Network Profiles:		
Facebook:		Twitter:
Instagram:		Google:
Other:		Other:
Child # 1 Social Network Profiles:		
Facebook:		Twitter:
Instagram:		Google:
Other:		Other:
Child # 2 Social Network Profiles:		
Facebook:		Twitter:
Instagram:		Google:
Other:		Other:
Other:		Other:

This partial template includes explicit details located during my investigation. I don't cite any sources, and I use this as an easy reference for account information that someone may use later. On some investigations, the Suspect Details section is several pages. After I have completed my Executive Summary and Suspect Details, I write the report narrative. The following are two small portions of how this report might appear. Note that I already possess screen captures of all online evidence, titled as explained in Chapter One.

Investigation Number: 2017-54887 Date: November 17, 2017
Investigator: Michael Bazzell Suspect: John Davis

On November 17, 2017, I was assigned an investigation into potential fraudulent medical claims made by John Davis, a former employee of INSERT COMPANY HERE. The following represents a detailed summary of my findings.

I located the Facebook profile of the suspect at facebook.com/JohnDavis9. I generated a screen capture of each section of this profile, as publicly visible to any user. These files were saved to disk, and titled as follows.

001-https__facebook.com_JohnDavis9 | 2017-11-17-10-15-11.pdf
002-https__facebook.com_JohnDavis9_photos | 2017-11-17-10-16-12.pdf
003-https__facebook.com_JohnDavis9_about | 2017-11-17-10-17-18.pdf
004-https__facebook.com_JohnDavis9_friends | 2017-11-17-10-19-31.pdf
005-https__facebook.com_JohnDavis9_events | 2017-11-17-10-22-11.pdf

Most notable within these captures is the "Photos" section identifying photos associated with hunting, including images with both Mr. Davis and his doctor within the same photo. This specific evidence is titled as follows.

006-https__facebook.com_photo.php?fbid=1828444224119932 | 2017-11-17-10-33-11.pdf

Note that I did not place a screen capture of this evidence within the report itself. There are two main reasons I do not place screen captures within printed report text. First, I believe it is better to keep the report concise and clutter-free. The client can easily view the evidence within the provided disc or drive. Second, it can be quite difficult to include an entire screen capture within an 8 ½ x 11 page. I would likely need to crop any images, which also means that I am redacting evidence. I would rather my client view the digital screen capture which displays the entire page. If I do want to include printed online screen captures, I will do so at the end of the report, and as a supplement. Also notice that after each set of screen captures, I summarized the value. In this example, the most beneficial evidence was a specific image. I have found that presenting the client with every possible detail results in an overwhelming report. I believe it is our job to tell the client what he or she may care about. After all, we are the analysts. Anyone can dump a bunch of screen shots. The true value is understanding why these captures are important. In the next example, I outline findings on Twitter.

I located the Twitter profile of the suspect's daughter, Kylie Davis, at twitter.com/kdavis722. I exported the most recent 3,200 posts (Tweets), and saved this as kdavis722.csv on the attached disc. I found the messages between Kylie Davis and Patricia Stevens (pstevens6655) of most interest. I isolated these messages with the following two queries.

from:kdavis722 to:pstevens6655
from:pstevens6655 to:kdavis722

Screen captures of these messages were saved as the following.

045-https__twitter.com_from:kdavis722 to:pstevens6655 | 2017-11-17-11-15-45.pdf
046-https__twitter.com_from:pstevens6655 to:kdavis722 | 2017-11-17-11-16-42.pdf

Of these messages, I found three references to the suspect and his doctor participating in a hunting trip. These specific references were cropped and saved as follows.

045a-Cropped Messages.pdf
046a-Cropped Messages.pdf

Note that I included details of the search technique, the specific evidence files, and information as to the importance of the content. I like to be as brief as possible. The digital screen captures provide all of the evidence necessary, and explicit detail of each capture is overkill. In most investigations, I have several pages of this type of narrative. Finally, I include a one-page Summary Report at the end. This also identifies future investigation needs and whether the incident is resolved. The following is a partial example.

This investigation was conducted with the hopes of identifying the participation of medical fraud by the suspect. I believe that this claim of fraud has been proven true. I advise continuous monitoring until the workman's comp claim is settled. Specifically, this investigation reveals the following as fact.

- Online evidence proves a personal association between the suspect and his doctor.
- Online pictorial evidence proves the suspect to have been physically fit enough to lift heavy objects within the time period of the disability claim.
- Online evidence proves the suspect to have been physically fit enough to run 5 kilometers within 28 minutes within the time period of the disability claim.
- Online evidence proves the suspect to be able to hunt in rugged conditions within the time period of the disability claim.

Note that I did not make any claims I could not prove, and I did not inject much opinion into the matter. I always try to keep the reports factual and unbiased. If I locate any online evidence that supports the suspect, I include it as well. When this happens, I make sure to emphasize this with digital screen captures and a brief summary of the content. Overall, I try to include the following with each report.

- Executive Summary: One-page synopsis of vital evidence.
- Suspect Details: Specific data such as all personal identifiers, user names, etc.
- Narrative Report: Detailed findings with references to digital evidence and summaries.
- Summary Report: One-page summary of facts and need for future work.
- Digital Evidence: A DVD or Drive that contains all screen captures and files.

Law enforcement can apply these same practices with one caveat. I believe that every criminal investigation should be conducted within a virtual machine. This could be the Buscador system mentioned in Chapter Two or a standard Windows or Linux system. At the end of the investigation, the entire machine should be exported as a single digital file and included with the digital evidence. I would also consider including the following paragraph within your narrative report.

> This entire investigation was conducted within a Linux virtual machine. This operating system was created on (insert date) and saved as a master copy. All security updates were applied at that time and no online investigation was conducted within this master copy. A clone of this system was created and titled (case number). This clone was used as the only operating system resource for the entire investigation. No other investigations were conducted within this clone. At the end of the investigation, this virtual machine was exported as (file name). This file can be used to recreate the entire investigation environment exactly as it appeared during the actual investigation.

This verbiage announces your competence to the prosecution and defense. It may stop some scrutiny toward your work during a trial or hearing. Ultimately, it shows that you conducted your investigation fairly with great concern for the integrity of your evidence. Additionally, this may make you stand out to your supervisors or the office of prosecution. I have found that consistent dedication to accurate reporting can go a long way toward your reputation and promotions.

This chapter has been overly simplified. Your reports may be extremely complex and contain dozens of pages. My goal here was to simply provide documentation considerations and their impact on the success of your investigation. Once you have developed a report template that works well for your investigations, recreating a report for each case will save time and energy. Everyone's reports are unique, and you should find the best way to explain your findings to an audience of any technical level.

Conclusion
What Now?

I hope the techniques presented have sparked your interest in finding new avenues of research and investigations. With patience and diligent effort, this book will be a helpful reference guide to assist in performing more accurate and efficient searches of open source intelligence. Permanently documenting these techniques on paper ~~may~~ will provide outdated content. Technology changes quickly and methods must adapt to keep up. Ten years from now, this book may be an amusing piece about how we once managed our online data. To keep up-to-date with the changes in various methods and OSINT data collection, please subscribe to my free monthly email newsletter at **IntelTechniques.com**. The chances are good that as you read this, new content has been posted about the very topic you are researching. The same website will allow you to access all of my current OSINT tools and links. Look for the "Tools" tab at the top of the menu. My free OSINT and Privacy Web Forum now has over 4,000 members that have posted over 10,000 comments. I learn something new there every day from the great minds that constantly share their research. I also offer an online OSINT and Privacy video training course that includes over 80 hours of HD videos, advanced resources, access to all of the APIs explained here, and a software pack of every application ready to use.

I am often asked my opinion about the future of OSINT. Occasionally, I am asked to advise intelligence collection companies during the creation of a new "all-in-one" OSINT solution. I decline these requests because the "easy" solutions are usually short-lived. Constant changes in technology, and automated data collection restrictions, make these commercial products less powerful over time. I do not believe these can ever replace the analytical brain. The truly valuable and powerful OSINT techniques are going to require manual work from an individual investigator. Your OSINT analysis skills cannot be replaced by a machine. Please go do good things with these methods.

Finally, remember that each of these investigation techniques could be used against you. When you find your own information exposed, take action to protect yourself. Please visit my other site Privacy-Training.com for hundreds of free resources including removal links, blog posts, and a weekly privacy and security podcast. Personal defense against OSINT is as important as offense.

Thank you for reading.

A special **THANK YOU** to Y.Varallo and the editor who insisted on remaining anonymous. You both make me appear smarter than I am. This book would be a mess without your input and guidance. I owe you both more credit than I can possibly give within this closing thought.

2Lingual, 77
360Social, 27
4chan, 189
4K Stogram, 374
Account Export, 178
Addresses, 240
Advanced Background, 228
Advangle, 82
Ahmia, 85
Aircraft Information, 358
Amazon, 202
Amazon AWS, 277
Analyze ID, 213, 328
Android Apps, 412
Android Emulation, 405
Antivirus, 1
APIs, 387
Archive.is, 73
Backpage, 196
BackTweets, 161
Baidu Cache, 72
Baidu Images, 287
Barcodes, 297
Bing, 70
Bing Advanced, 81
Bing Cache, 72
Bing Images, 71, 285
Bing IP, 340
Bing Maps, 263
Bing Operators, 70
Bing Translator, 76
Bing Videos, 313
Birth Records, 355
BleachBit, 382
Board Reader, 194
Bootable USB, 51
Broadcastify, 440
Built With, 330
Bulk Downloader, 10
Burner, 253
Buscador, 33
Business Records, 354
Caller ID, 244, 397
Caller ID Service, 245
Caller ID Test, 250
Camera Trace, 295
Campaign Contributions, 358
CamStudio, 375
Carrot2, 83
CCleaner, 3, 382
Check User Names, 218
Chrome, 25
Citizen Band, 433
City Vibe, 197
ClamAV, 2
Classmates, 235
Compromised, 211, 221, 400
Contact Exploitation, 175, 415
Copy All Links, 22
Coral, 73
County Records, 353
Court Records, 353
Craigslist, 194, 198, 256
Credit Card Data, 361
Creepy, 372
Criminal Data, 358
Cubib, 232
Custom Search Engines, 65
Dating Websites, 191
Death Records, 355
Descartes Labs, 264
DNS Trails, 213, 325
Document Metadata, 43
Documentation, 451
Documents Search, 275
Documents, 275
Domain Analytics, 328
Domain Big Data, 213
Domain Crawler, 329
Domain History, 325
Domain Hosting View, 381
Domain Names, 321
Domain Registration, 321
Download Helper, 9
Duck Duck Go, 89
eBay, 200
Email Addresses, 207
Email Assumptions, 208
Email Hunter, 207, 335
Email Permutator, 209
Email Spoofing, 350
Email Validation, 210
Erotic Review, 197
Escort Ads, 197
Escort Reviews, 196
Etcher, 54
Exalead, 84
Exif Tool, 373
Exif Viewer, 17
ExtractFace, 380
EyeWitness, 45
Facebook, 95
Facebook:
  Accounts, 133
  Common Friends, 113
  Common Results, 111
  Email, 130
  Events, 116
  Friends, 108, 113, 132
  ID, 104, 114
  Photos, 131
  Posts, 102
  Profiles, 102, 114
  Search, 97, 127
  Telephone, 127
  Videos, 118, 313
Fake Followers, 160
FakeSpot, 203
Family Radios, 432
Family Tree Now, 231
Fast People, 230, 251
FFmpeg, 40, 364
File Mare, 91
Find My Snap, 220
Find People, 231
Firefox, 5
Firefox Add-ons, 8
Firefox Profile, 23
FireShot, 11
Flickr, 292, 396
Flippity, 201
FOCA, 378
FollerMe, 163
Follow That Page, 327
Followerwonk, 137, 157
Foreign Search, 77
Forensically, 299
Forums, 194
Foto Forensics, 298
FTP Search, 89
Full Contact, 392

Genymotion, 406
Gift registries, 283
Global File Search, 89
Gmail, 96
Gmail Contacts, 176
GoofBid, 201
Google:
  Search, 57
  Advanced, 81
  Alerts, 69
  Blogs, 81
  Cache, 71
  Docs, 276
  Earth, 370
  Groups, 78
  Images, 71, 285
  Input, 77
  Maps, 261
  News, 79
  Newspapers, 79
  Operators, 57
  Scholar, 82
  Takeout, 178
  Tools, 63
  Translator, 20, 76
  Videos, 313
Google+, 170
Government Records, 353
Gravatar, 212
Hacked-Emails, 211, 402
Hacker News, 190
Harvester, 45
Have I Been Pwned, 211, 401
Here Maps, 263
Hootsuite, 157
HTTPS Everywhere, 17
HTTrack, 374
Hunchly, 28
Image Manipulation, 297
Image Raider, 290
Image Search Options, 21
Infobel, 259
Instagram, 96, 165
Instagram Accounts, 167
Instagram Followers, 168
Instagram Images, 166
Instalooter, 48
Intelius, 229
IntelTechniques Tools:
  Search Engines, 93
  Facebook, 124
  Twitter, 145
  Instagram, 169
  LinkedIn, 173
  Communities, 204
  Email, 214
  User Names, 221
  People Search, 233
  Telephone Numbers, 255
  Maps, 267
  Documents, 279
  Paste Sites, 283
  Reverse Images, 289
  YouTube, 308
  Reverse Video, 312
  Domain Names, 337
  IP Addresses, 345
International Networks, 179
International Search, 86
Internet Archive, 315
IP Addresses, 339
IP Logging, 346
IP Net Info, 382
IPLocation, 339
iSEEK, 83
Izitru, 298
JavaScript Bookmarklets, 24
Jeffrey's Exif Viewer, 293
JPEG Snoop, 373
Karma Decay, 290
KeePassXC, 385
Keyword Tool, 83
KnowEm, 217
Land Viewer, 266
LibreOffice, 51
Lightshot, 376
LinkedIn, 171
Live Video Streams, 317
Loyalty Cards, 259
Malware Bytes, 3
Mamont, 91
Many Contacts, 212
Map Box, 263
Mapillary, 264
Maps, 261
Marine Channels, 434
Marine Traffic, 357
Meetup, 190
Melissa Data, 230
MentionMapp, 161
Metadata, 281, 293, 369
Metadata Toolkit, 50
Metagoofil, 43
Military Records, 360
Million Short, 85
Million Tall, 85
MJSONViewer, 18
MySpace, 175
Name Checkr, 218
Name Chk, 217
NameVine, 218
Napalm FTP, 91
Nerdy Data, 91, 329
Newspaper Archive, 79
Newspaper Comments, 195
Next Caller, 249
Nimbus, 12
Nuwber, 231
OfferUp, 202
Onion Link, 85
Online Communities, 183
Online Prostitution, 196
Online Translator, 76
Open CNAM, 245, 247
Open Street Cam, 264
OSINT Workflow, 443
Paste Sites, 283
Pastebin, 283
Peek You, 219
People Search, 227
People Search Now, 231
Periscope, 318
Photographs, 285
Pictriev, 290
Pinterest, 203
Pipl, 212, 218, 227, 251, 389
Presentations, 278
Private Profiles, 175
Prophet, 26
PubDB, 329
Quanki, 232
Qwant, 93
Qzone, 180
Radio Frequencies, 431

Radio Reference, 437
Recon-ng, 419
Recruit'em, 173
Recuva, 385
Redaris, 230
Reddit, 183
Reddit Domains, 335
Reddit Images, 187
Reddit Sub-Reddits, 187
Renren, 180
Rental Vehicles, 282
Resumes, 236
Resurrect Pages, 21
Reverse Image Search, 285
Reverse Video Search, 308
Robots.txt, 330
Roofus, 53
Satellite View, 261
Screenshots, 326
Scribble Maps, 273
Scribd, 279
Script Blockers, 13
Search Engines, 57
Search Tempest, 201
Searx, 84
SEO Spider, 381
SEO Tools, 332, 334
Service Objects, 256, 398
Shared Count, 334
Shodan, 343
Shortened URLs, 336
Skype, 220
Sly Dial, 258
SmartDeblur, 377
Snapchat, 175
Social Mention, 177
Social Searcher, 177
Software Radio, 431
SpiderFoot, 49
Spokeo, 228
Spy Dialer, 258
Spy On Web, 328
Start Page, 89
Status People, 160
Stolen Camera Finder, 294
Street View, 262, 264
Stumble Upon, 203
Subdomains, 47, 330
Talk Walker, 69
Taringa, 181
Telephone Carries, 243
Telephone Numbers, 243
Telephone Search, 253
Terra Server, 266
That's Them, 212, 228, 341
Tinder, 192
TinEye, 287
Tinfoleak, 48, 162
Topix, 204
Tor Browser, 29
Tor Scan, 86
Tor Search, 85, 86
Tor2Web, 85
TowerData, 399
Trendsmap, 161
True Caller, 255
True People, 231, 252
Truth Finder, 233
Tumblr, 174
TV News Archive, 316
Tweet Deck, 156
TweetBeaver, 147
TweetTopic, 163
Twiangulate, 157
Twilio, 245
Twitonomy, 161
Twitter:
  Search, 96, 135
  Audit, 160
  Bios, 144
  Contacts, 176
  Deletions, 141
  Directory, 137
  Export, 47
  Images, 291
  Locations, 150
  Media, 141
  Operators, 138
  Search, 135
uBlock Origin, 13
User Agent Switcher, 19
User Names, 217
User Search, 218
User Sherlock, 220
Vehicle, 356
Vehicle Registration, 357
VeraCrypt, 383
Video Closed Captions, 316
Video Download, 42,306,319,367
Video Utilities, 40, 364
Videos, 303
ViewDNS, 322, 340
VIN Search, 356
Vine, 317
Virtual Currencies, 362
Virtual Device Cloning, 416
Virtual Device Export, 417
Virtual Machines, 33
Virtual Private Network, 30
Virtual Snapshots, 37
VirtualBox, 35
Virus Total, 335
Visual Ping, 327
Visual Site Mapper, 335
VK, 179
VMWare, 36
Voat, 189
Voter Registration, 241, 361
Wayback Machine, 74
Web SDR, 441
WebMii, 232
WhoCalld, 245
Whoismind, 213, 326
Whoisology, 324
Whoxy, 213
Wigle, 342
WikiLeaks, 280
Wireless Microphones, 436
Wireless Monitors, 435
Yahoo, 96
Yahoo Groups, 78
Yandex, 87
Yandex Cache, 72
Yandex Images, 287
Yandex Operators, 87
Yasni, 229
YouTube, 303, 307
YouTube Comments, 307
YouTube Restrictions, 304
YouTube-DL, 42
Zaba Search, 230
Zillow, 241
Zoom Earth, 263
ZoomEye, 344



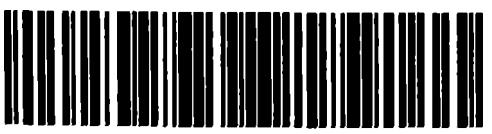
Inteltechniques.com

Online OSINT Training Videos

Thank you for purchasing a copy of this book. Use the following website to receive 25% off your purchase of monthly or yearly access to over 195 online OSINT training videos.

inteltechniques.com/25

Printed in Great
Britain
by Amazon



Sixth Edition Sheds New Light on Open Source Intelligence Collection and Analysis

Author Michael Bazzell has been well known in government circles for his ability to locate personal information about any target through Open Source Intelligence (OSINT). In *Open Source Intelligence Techniques: Resources for Searching and Analyzing Online Information*, he shares his methods in great detail. Each step of his process is explained throughout twenty-five chapters of specialized websites, software solutions, and creative search techniques. Over 250 resources are identified with narrative tutorials and screen captures. This book will serve as a reference guide for anyone that is responsible for the collection of online content. It is written in a hands-on style that encourages the reader to execute the tutorials as they go. The search techniques offered will inspire analysts to "think outside the box" when scouring the internet for personal information. Much of the content of this book has never been discussed in any publication. Always thinking like a hacker, the author has identified new ways to use various technologies for an unintended purpose. This book will greatly improve anyone's online investigative skills. Among other techniques, you will learn how to locate:

Hidden Social Network Content
Cell Phone Subscriber Information
Deleted Websites & Posts
Missing Facebook Profile Data
Full Twitter Account Data
Alias Social Network Profiles
Free Investigative Software
Useful Browser Extensions
Alternative Search Engine Results
Website Owner Information
Photo GPS & Metadata
Live Streaming Social Content
Social Content by Location
IP Addresses of Users
Additional User Accounts
Sensitive Documents & Photos

Private Email Addresses
Duplicate Video Posts
Mobile App Network Data
Unlisted Addresses & #s
Public Government Records
Document Metadata
Rental Vehicle Contracts
Online Criminal Activity
Personal Radio Communications
Compromised Email Information
Automated Collection Solutions
Linux Investigative Programs
Dark Web Content (Tor)
Restricted YouTube Content
Hidden Website Details
Vehicle Registration Details

Michael Bazzell spent 18 years as a government computer crime investigator. During the majority of that time, he was assigned to the FBI's Cyber Crimes Task Force where he focused on open source intelligence (OSINT) collection and analysis. He has trained thousands of individuals employed by state and federal agencies, as well as the private sector, in the use of his OSINT investigation techniques. He is also the author of *Hiding from the Internet*. His books are used by numerous government agencies as training manuals for intelligence gathering and proper securing of personal information.

$ 44.99 US
£ 29.99 UK
€ 37.99 EU

ISBN 9781984201577

INTELTECHNIQUES.COM