How to conduct effective Open Source Investigations online

Vytenis Benetis | 13 Oct 2020 | ITU 2020 Global CyberDrill
UNOCT/UNCCT Consultant
Why do we need to become good online investigators?
Malicious use of Internet

- Malicious software
- (D)DoS
- Phishing
- Radicalization
- Propaganda
- Psychological warfare, denial and deception
- Intelligence gathering
- Cyberattacks and information operations
- Money laundering
- ....
Size of internet

**DIGITAL AROUND THE WORLD IN 2020**

**Total Population**: 7.75 Billion, Urbanisation: 55%

**Unique Mobile Phone Users**: 5.19 Billion, Penetration: 67%

**Internet Users**: 4.54 Billion, Penetration: 59%

**Active Social Media Users**: 3.80 Billion, Penetration: 49%

*Sources: Population - United Nations, Local Government Bodies; Mobile - GSMA Intelligence, Internet; ITU; GlobalWebIndex; GSMA Intelligence; Local Telecoms Regulatory Authorities and Government Bodies; AT&T, Kaspersky, Analysys, Social Media Platforms, User Service Advertising Tools, Company Announcements and Earnings Reports; Caterpillar, Kaspersky, Analysys. All latest available data in January 2020. For comparability, we have sourced and base changes.*
Technological advancements

- Infrastructure (networks, devices)
- Computing power
- Content digitization
- AI (touches every aspect)
- Mobility
- Interface and sensorics
- Cybercurrency
- IoT
- ...
“On the Internet, nobody knows you’re a dog.”
Future Challenge of Internet

• The Internet will continue to fragment across regional, national and ideological lines, making the job of intelligence collection harder

• Our reluctance to embrace OSINT literacies to enhance intelligence collection will hurt our ability to
“Meta” aspects of intelligence
What is intelligence?

- Business of managing uncertainty
- Informant to decision making
- A form of knowledge, a form of organization and a form of activity (*Sherman Kent*)
Decision window

Better intelligence process

Decision flexibility

Information quality

Decision window

Time
Important Role of OSINT

“We have no need of spies. We have the Times of London.” Tsar Nicholas II (1818-1881)

“A proper analysis of the intelligence obtainable by these overt, normal, and aboveboard means would supply us with over 80 percent, I should estimate, of the information required for the guidance of our national policy.” Allen Dulles, Head of the CIA (1893 – 1969)

“More can be deduced by an intelligent study of public sources than by any number of ‘reliable’ but unintelligent ‘agents’ listening at keyholes or swapping drinks at bars. Hugh Trevor-Roper (1914 – 2003)
Important Role of OSINT
Legal and ethical considerations

• Information is your basic tool of work

• Information is subject to laws, which are never very clear

• The more information you collect, greater the risk you will break the law

• Technology amplifies the risks as well as the mistakes we make

• Thus, information should always be handled with care
Legal and ethical considerations

- Copyright
- Defamation
- Privacy
- Legislation: local vs. global
- Jurisdiction
- Social attitudes and values
Avoid legal headaches

• Educate yourself

• Question your motives

• If in doubt, ask a lawyer
Avoid legal headaches

• Educate yourself
• Question your motives
• If in doubt, ask a lawyer
OSINT frameworks
Raison d'être for structured approach

• Have a map to navigate the “chaos”
• Produce an audit trail
• Pave the way for improvement
OSINT frameworks

• Research and monitoring
  • Keywords index
  • Source management
  • Risks/Threats early warning setup
  • Automated collection

• Investigation
  • Technical, procedural and analytical tools to scope one’s target
  • Hypothesis building / evidence collection
  • Process cannot be automated
Requirements
Learn to ask questions

Effective intelligence begins by addressing the following questions:

- What do we need to know?
- Why do we need to know it?
- Who might have the information we need?
- How should we perform the research?
- What will we do with the results?
- Does the effort justify the cost?
Requirements planning

What needs to be achieved:

- Clarified requirements
- Clarified goals and priorities
- Resources allocated
- Realistic timeline defined
- Relevant stakeholders engaged
Requirements planning

How to do it?

• Structured brainstorming techniques (Starbursting, KIT)
• Requirements frameworks and templates
• Technology aid (notetaking, mind mapping, researching)
Online Security
Setting up work environment

- Comprehensive management of all elements

Diagram:

- Hardware
- Internet access
- Software
- Cyber persona

+ Operating Process
Setting up work environment

• Hardware
  • Acquired with cash, wiped out, installed with bare bones
  • Disabled camera / mic/sound / location / Bluetooth

• Internet connection
  • Cash prepaid anonymous SIM card with prepaid data plan
  • Cash paid basic mobile device that tethers internet
  • Public hotspot (but be careful!)
  • Always privacy software layer on top
Setting up work environment

- Software tools to hide your identity
  - VPN (for ex. Proton VPN) - kill switch is a must!
  - Virtual Machine (VirtualBox + some Linux version for ex. Parrot OS)
  - TOR / Various privacy-oriented browsers

- Software tools to collect / automate evidence gathering
  - Hunchly ($)
  - Screenshot grabbers (Greenshort / Fireshot / Lightshot)
  - Maltego Casefile
  - Evidence collection templates
Setting up work environment

• Cyber personas
  • If required, created using secure email and prepaid sim card (for verification)
  • Meticulous buildup of cyber personas: consistency of the narrative with time zone, language settings, browser agent ID, VPN
Evidence collection process

• Operating process
  • Threat model for privacy / security
  • Lean Intelligence / People Search model (or others)
  • Templates to structure your information (for ex. Person Profile)
  • Selection of proper skillsets required: monitoring / investigation
  • Success criteria and measurement of progress against it
Online Security

• Who Are We Protecting Ourselves From?
  • Our targets
  • Hackers and cybercriminals
  • Internet Service Provider (ISP)
  • Advertisers and corporations
  • Employers
  • Other governments
Investigation techniques: People Search
People Search: Introduction
People search opportunities

Easier to obtain:

• The growth of available information
• Our interactions with the web
• The popularity of social networks
• Commercial value of people’s data

Face challenges:

• Not everyone has a digital footprint
• Information is dispersed
• Name match
• Intense time and labor required
People search process

- Identify names, usernames, aliases
- Identify social media accounts
- Identify contact details

- Run general search engine query
- Run custom search query
- Search databases, networks, etc.

- Requirements & Intel gaps
- Validate results
- Compile report
### People search process

<table>
<thead>
<tr>
<th>Category</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identifying phrases</td>
<td>Names, aliases, usernames, titles, etc.</td>
</tr>
<tr>
<td>Basic Information</td>
<td>Age, gender, ethnicity, nationality, spoken language, etc.</td>
</tr>
<tr>
<td>Contact Details</td>
<td>Telephone number, e-mail, Skype handle, etc.</td>
</tr>
<tr>
<td>Residence</td>
<td>Country of residence, current / past home address, profile of neighbourhood, etc.</td>
</tr>
<tr>
<td>Family</td>
<td>Marital status, spouse / partner, children, parents, siblings, cousins, etc.</td>
</tr>
<tr>
<td>Work</td>
<td>Employment status, current / past employer(s), office colleagues, etc.</td>
</tr>
<tr>
<td>Education</td>
<td>Level of education, attended educational institutions, classmates, studied subjects, etc.</td>
</tr>
<tr>
<td>Friends</td>
<td>Best friend(s), other friends, colleagues, acquaintances, etc.</td>
</tr>
<tr>
<td>Hobby &amp; Interests</td>
<td>Key hobbies, online interests, listened music, read books, watched movies, etc.</td>
</tr>
<tr>
<td>Views &amp; Opinions</td>
<td>Religion views, political views, likes / dislikes, etc.</td>
</tr>
</tbody>
</table>
People search process

- Decide how to organize / collate data
- Don’t break the law
- Identify formal names
- Identify titles and honorifics
- Identify the target’s social media profiles
- Identify the target’s contact details
- Identify the target’s usernames
- Identify the target’s locations
- Identify the target’s affiliations
Usernames

• Run usernames through discovery tools
• Verify as tools may not be perfect
• Check variations
• More popular tools:
  • Knowem: http://knowem.com
  • NameChk: www.namechk.com
  • More tools: https://lnkd.in/d_4K9HG
Email search

- Usernames often associate with emails
- Run Google queries / setup Google Alerts
- Check breached data (https://haveibeenpwned.com etc)
- Find private email address (constructs and guesses, socmint)
- Find professional email address (www.hunter.io etc)
- Run email validator (www.email-validator.net etc)
- Reverse email checks (www.pipl.com etc)
- Check email provider for business emails (www.mxtoolbox.com etc)
- Check blacklists (www.mxtoolbox.com etc)
Phone numbers

- Start with phone directories ([www.numberway.com](http://www.numberway.com) etc)
- Run Google queries / setup Google Alerts
- Check reverse phone lookup ([www.truecaller.com](http://www.truecaller.com) etc)
- Check Skype, Social media accounts
Google Hacks for People Search

• Master Search Engine operators
  Ex: “username” site:facebook.com inurl:photos; “username” inurl:profile

• Check online spaces (websites, blogs, wikis etc)
  Ex: site:wix.com “username”

• Check Q&A sites (quora, stackexchange, answers etc)
  Ex: “username” site:stackexchange.com

• Check user groups
  Ex: “username” site:groups.google.com

• Search document repositories (Docs, Aws, OneDrive, Slideshare etc)
<table>
<thead>
<tr>
<th>Operator</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR</td>
<td>Used to find synonymous or related content (write in uppercase)</td>
</tr>
<tr>
<td>-</td>
<td>The NOT operator hides / excludes unwanted keywords</td>
</tr>
<tr>
<td>&quot;Quote marks&quot;</td>
<td>Returns the exact combination of words between the quote marks</td>
</tr>
<tr>
<td>filetype:</td>
<td>Reduces results to specific file types</td>
</tr>
<tr>
<td>related:</td>
<td>Will help you identify web pages similar to your specified site</td>
</tr>
<tr>
<td>site:</td>
<td>Results limited to a specific website or domain</td>
</tr>
<tr>
<td>intitle: / allintitle:</td>
<td>Results limited to those pages with the keywords in the title</td>
</tr>
<tr>
<td>inurl: / allinurl:</td>
<td>Results limited to those sites with the keyword in the URL</td>
</tr>
<tr>
<td>intext: / allintext:</td>
<td>The query is limited to the text of a page only</td>
</tr>
<tr>
<td>*</td>
<td>Use the wildcard operator for spelling and phrase variations variations</td>
</tr>
<tr>
<td>..</td>
<td>Use the range operator to search for a range of numbers</td>
</tr>
</tbody>
</table>
Google Hacks for People Search

• Search for CVs (ex: “Name” “CV” inurl:resume OR intitle:resume

• Check dating sites (username + “site” operator)

• Check online marketplaces (username + site or inurl; ex: alibaba.com)

• Education history (site: + domain or education institution)

• Validate credentials through complex queries (ex: intitle:“TARGET NAME”
inurl:speaker OR inurl:speakers OR inurl:author OR inurl:authors OR
inurl:instructor OR inurl:instructors OR inurl:expert OR inurl:experts
Effective search model

Source knowledge

Search Engine

Keywords

Operators

Smart queries
Working with images
Reverse Image search

- Reverse search for image
  - Google: www.google.com/images
  - Tin Eye: www.tineye.com
  - RootAbout: http://rootabout.com
  - Yandex: https://yandex.com/images
  - Baidu: http://image.baidu.com

- Analyze metadata (exif + content on the host site)
  - Jeffrey Friedl’s Exif Viewer: http://exif.regex.info/exif.cgi
  - Megapicz: http://metapicz.com/#landing

- Forensic analysis
  - Fotoforensics http://fotoforensics.com
  - Image forensic https://www.imageforensic.org
WEBINT
## Digital data hierarchy

<table>
<thead>
<tr>
<th>Individual Data</th>
<th>Organisational Data</th>
<th>Network Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Key personnel</td>
<td>• Business locations</td>
<td>• IP Data</td>
</tr>
<tr>
<td>• Contact details</td>
<td>• Company addresses</td>
<td>• Internal domain names</td>
</tr>
<tr>
<td>• Email addresses</td>
<td>• Phone numbers</td>
<td>• Name servers</td>
</tr>
<tr>
<td>• Email conventions</td>
<td>• Security policies</td>
<td>• Email servers</td>
</tr>
<tr>
<td>• Phone numbers</td>
<td>• Web service providers</td>
<td>• Web technologies</td>
</tr>
<tr>
<td></td>
<td>• Social media assets</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• System technologies</td>
</tr>
</tbody>
</table>
WEBINT Toolkits

- Central Ops: [http://centralops.net](http://centralops.net)
- Domain Big Data: [https://domainbigdata.com](https://domainbigdata.com)
- Domain Tools: [http://research.domaintools.com](http://research.domaintools.com)
- Hacker Target: [https://hackertarget.com/ip-tools](https://hackertarget.com/ip-tools)
- Kloth: [www.kloth.net/services](http://www.kloth.net/services)
- Network Tools: [www.network-tools.com](http://www.network-tools.com)
- MX Toolbox: [www.mxtoolbox.com](http://www.mxtoolbox.com)
- You Get Signal: [www.yougetsignal.com](http://www.yougetsignal.com)
Investigate websites/domains

1. Identify the Whois Data
2. Reverse IP / DNS Lookup
3. Check a Website’s IP History
4. Analyze Hosts [DNS Dumpster / Yougetsignal]
5. Investigate Subdomains [Security Trails etc]
6. Identify Other Services Running on a particular IP
7. IP Mapping
Investigate websites

8. Examine Digital Certificates
9. Study the Website’s HTML Data (traces, tracking IDs)
10. Check Robots.txt / Sitemap
11. Check site’s security and reputation [VirusTotal etc]
12. Site technologies
13. Backlinks using SEM tools [Linkminer, Semrush etc]
Tor Investigations
Tor Investigations Framework

**SURFACE WEB**
- Find .onion sites of interest
- Search keywords from dark web including .onion addresses
- Setup alerts on the keywords
- Run reverse image search on images from dark web

**DARK WEB**
- Visit .onion sites and collect seeds for further investigation
- Use Dark web search engines to find further .onion sites

Collate and analyze data; Identify intel gaps; Formulate queries.

**Crawler service** (3rd party or self-host ex. TorBot, TorCrawl)
Setup your workplace

- Access the Tor network using a “clean” device / sensors disabled
- Run Tor engine inside a virtual machine
- Choose carefully your entry network node
- Setup data capturing solution
Operational security tips

• Before opening the Tor browser, close all other software running on your system and disable any plugins in the browser.

• Generate “New Identity” or “New Tor Circuit” every time you access a new .onion link.

• Do not download any content unless necessary.

• Use sock puppet accounts.
Tor Investigations

• Take note...

  – The Tor network’s popularity has been rather steady over last few years

  – Tor is a high priority target for security services, which are busy identifying and exploiting vulnerabilities in the browser

  – For this reason, maintaining one’s security while using Tor is now more important than ever
Cryptocurrency Investigations
Cryptocurrency investigation framework

1. Find and collect seeds
2. Monitor seeds
3. Select the toolbox
4. Analyze evidence
5. Resolve to suspect
Cryptocurrency investigation resources

Books, articles

Web tools

Specialized products

- Address Checker
- Aware Online: Cryptocurrency Search Tool
- Bitcoin Abuse
- Bitcoin Block Resources
- Bitcoin.org
- Bitcoin WhoisWho
- BitInfoCharts
- Bitnodes
- bitonic
- BitRef
- Blockchain.com
- Blockchain
- Blockchain Engine
- Blockchain Explorers
- Blockhars
- BlockCypher
- Blockonomics
- Blocks.Press
- BTC.com
- BTCScans
- Bytc.io

- Chainalysis
- CipherTrace
- DataWalk
- Elliptic
Tracking Bitcoin transactions

• After finding a dark web website or a content, note any cryptocurrency addresses and other related identifiers you can find

• Run them through surface and dark web-based cryptocurrency explorers:
  
  • Wallet Explorer: [www.walletexplorer.com](http://www.walletexplorer.com)
  
  • Blockchain Block Explorer: [www.blockchain.com/explorer](http://www.blockchain.com/explorer)
  
  • Bitcoin WhosWho: [https://bitcoinwhoswho.com](https://bitcoinwhoswho.com)
  
  • Bitcoin Abuse Database: [www.bitcoinabuse.com](http://www.bitcoinabuse.com)