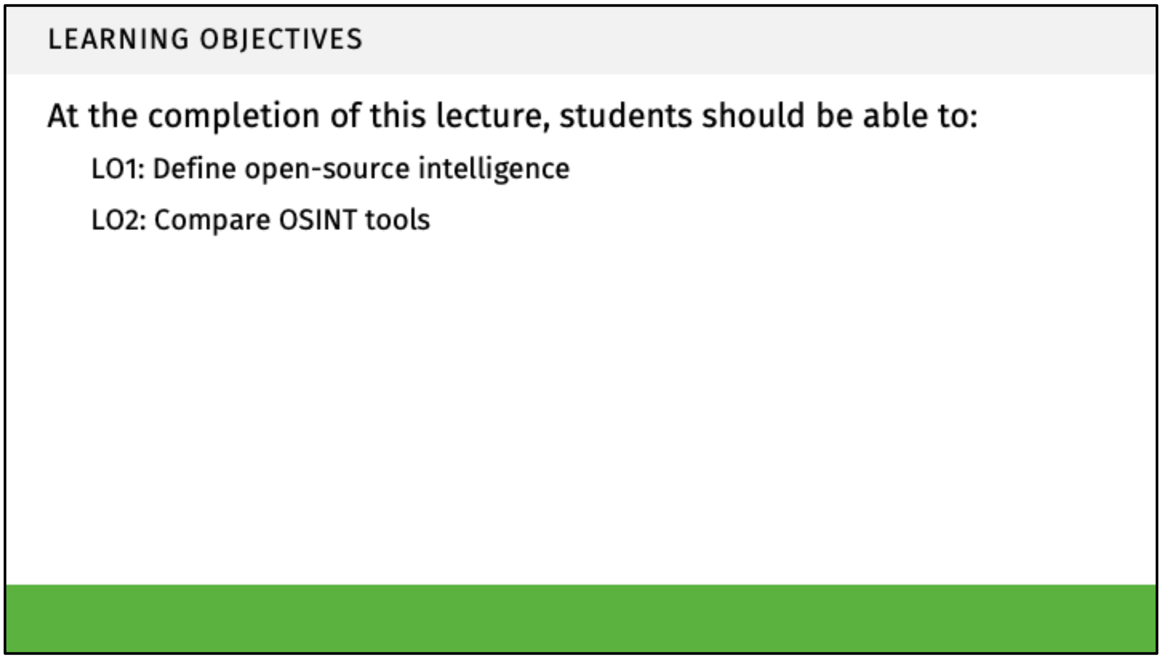
APP100

Intelligence Gathering

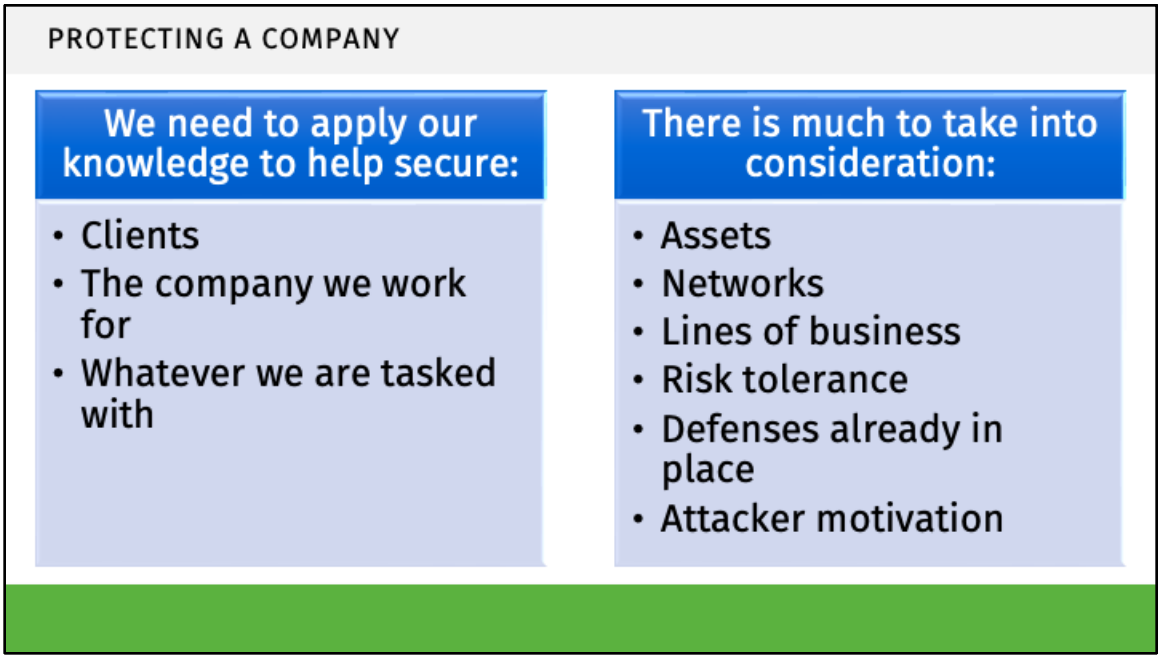


LEARNING OBJECTIVES

At the completion of this lecture, students should be able to:

LO1: Define open-source intelligence

LO2: Compare OSINT tools



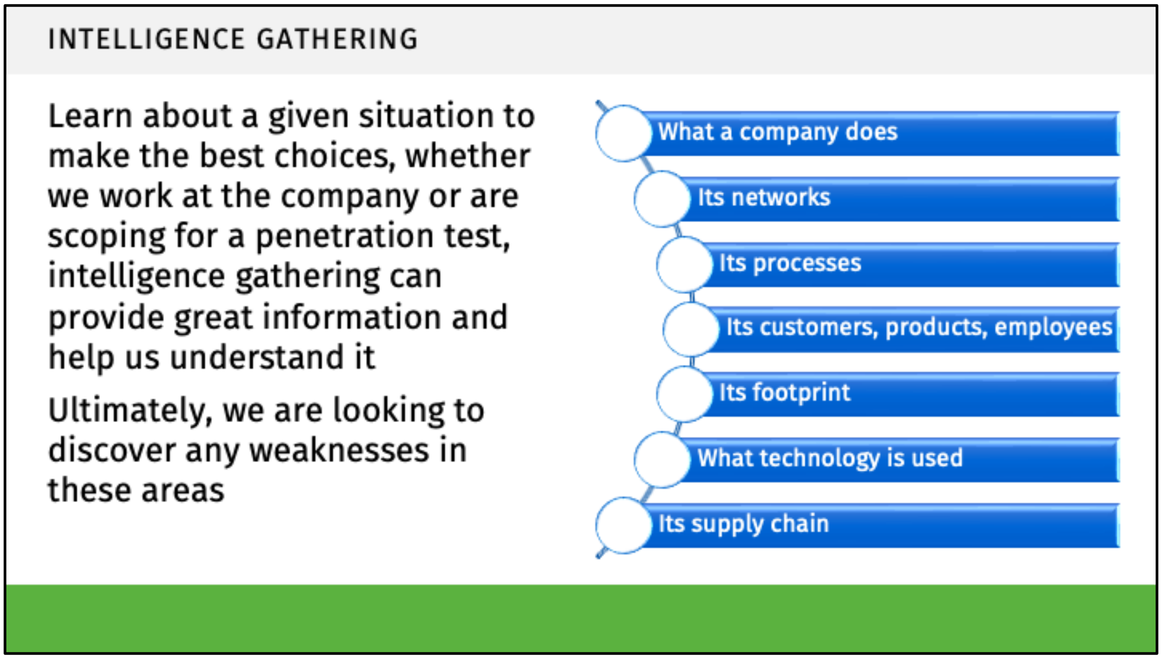
PROTECTING A COMPANY

We need to apply our knowledge to help secure:

- Clients
- The company we work for
- Whatever we are tasked with

There is much to take into consideration:

- Assets
- Networks
- Lines of business
- Risk tolerance
- Defenses already in place
- Attacker motivation

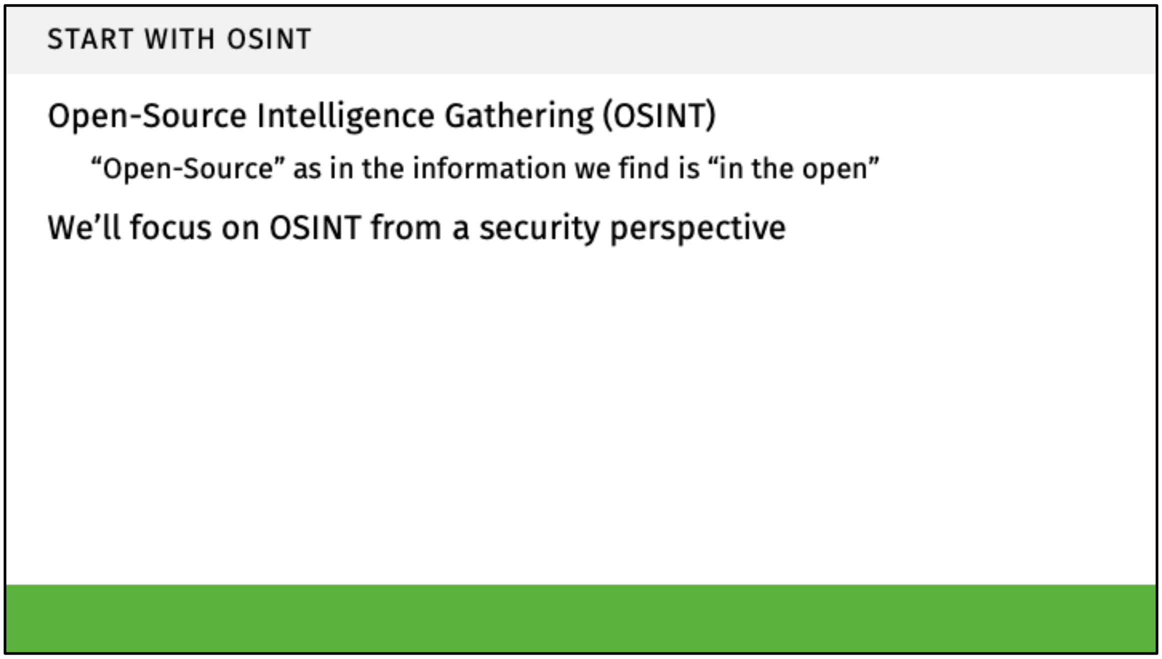


INTELLIGENCE GATHERING

Learn about a given situation to make the best choices, whether we work at the company or are scoping for a penetration test, intelligence gathering can provide great information and help us understand it

Ultimately, we are looking to discover any weaknesses in these areas

- What a company does
- Its networks
- Its processes
- Its customers, products, employees
- Its footprint
- What technology is used
- Its supply chain



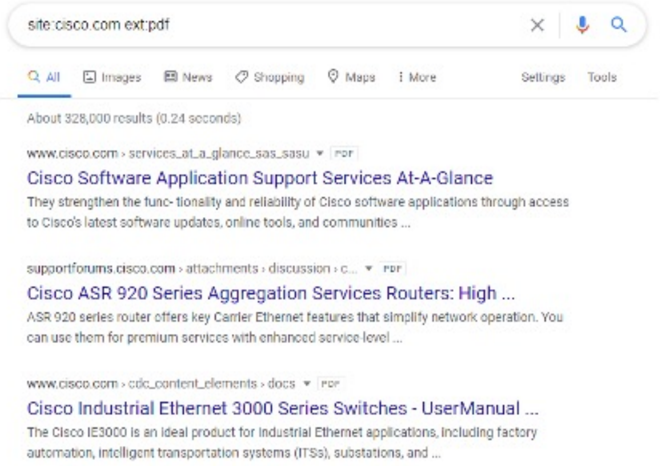
START WITH OSINT

Open-Source Intelligence Gathering (OSINT)

- “Open-Source” as in the information we find is “in the open”

We’ll focus on OSINT from a security perspective

WHAT IS THIS?



site:cisco.com ext:pdf

About 328,000 results (0.24 seconds)

www.cisco.com › services_at_a_glance_sas_sasu ▾ PDF
Cisco Software Application Support Services At-A-Glance
They strengthen the func- tionality and reliability of Cisco software applications through access to Cisco's latest software updates, online tools, and communities ...

supportforums.cisco.com › attachments › discussion › c... ▾ PDF
Cisco ASR 920 Series Aggregation Services Routers: High ...
ASR 920 series router offers key Carrier Ethernet features that simplify network operation. You can use them for premium services with enhanced service level ...

www.cisco.com › cdc_content_elements › docs ▾ PDF
Cisco Industrial Ethernet 3000 Series Switches - UserManual ...
The Cisco IE3000 is an ideal product for industrial Ethernet applications, including factory automation, intelligent transportation systems (ITSs), substations, and ...

JUST A GOOGLE SEARCH RESULTS PAGE, RIGHT?

We used the **site** operator

We used the **ext** operator

We EXPLICITLY told Google to ONLY show us results from ONE domain: cisco.com and ONLY show us PDFs from that domain

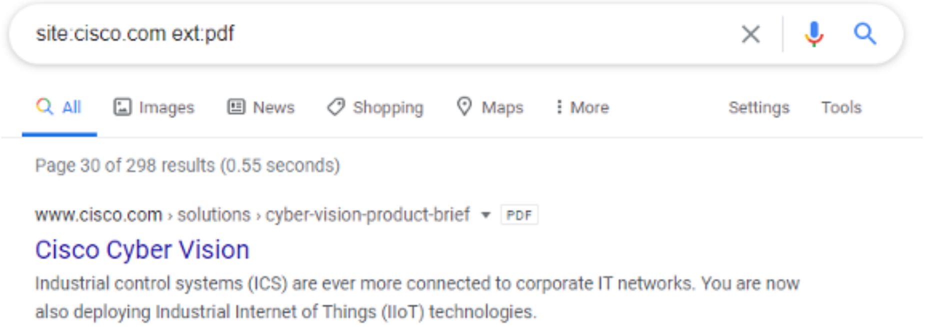
So what?

INFORMATION GATHERING

Now we have a listing of PDFs

Note the original query said there were 328,000 results (yeah, right)

Go to page 30 of Google results and we see there are 298

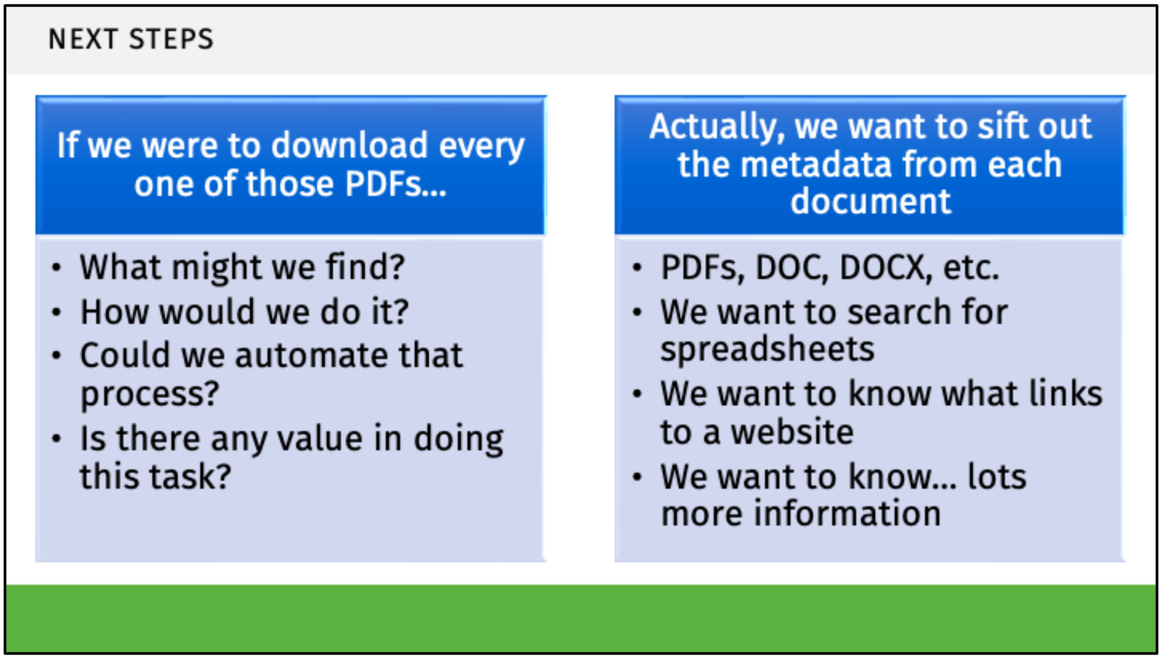


Page 30 of 298 results (0.55 seconds)

www.cisco.com › solutions › cyber-vision-product-brief ▾ PDF

Cisco Cyber Vision

Industrial control systems (ICS) are ever more connected to corporate IT networks. You are now also deploying Industrial Internet of Things (IIoT) technologies.



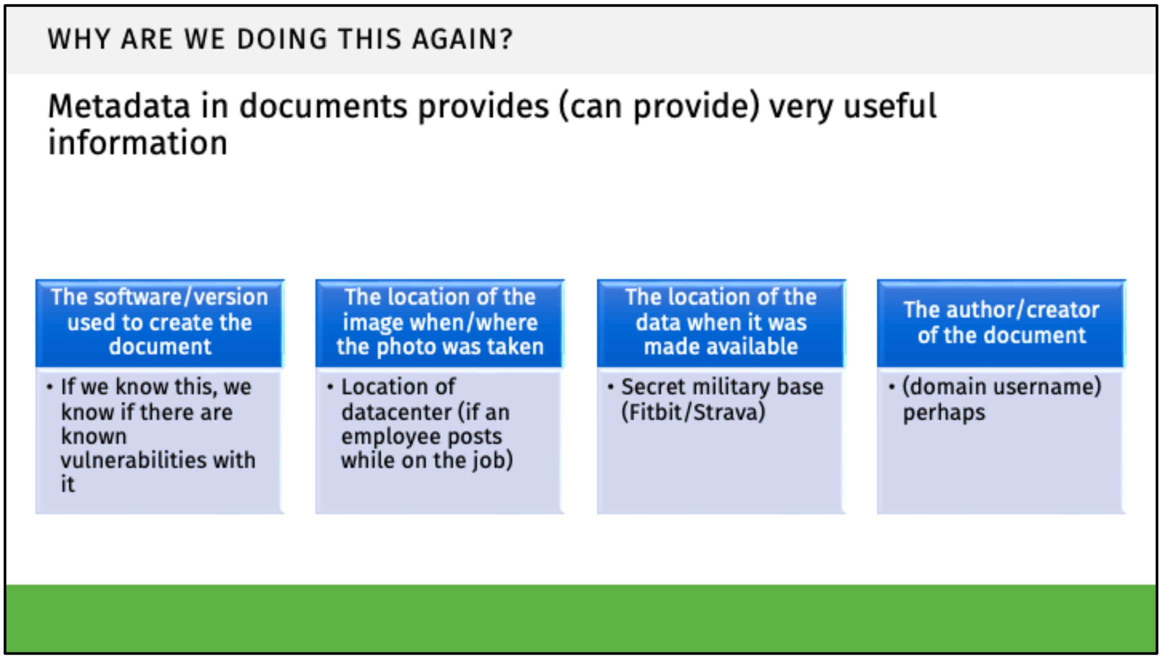
NEXT STEPS

If we were to download every one of those PDFs…

- What might we find?
- How would we do it?
- Could we automate that process?
- Is there any value in doing this task?

Actually, we want to sift out the metadata from each document

- PDFs, DOC, DOCX, etc.
- We want to search for spreadsheets
- We want to know what links to a website
- We want to know… lots more information



WHY ARE WE DOING THIS AGAIN?

Metadata in documents provides (can provide) very useful information

The software/version used to create the document

- If we know this, we know if there are known vulnerabilities with it

The location of the image when/where the photo was taken

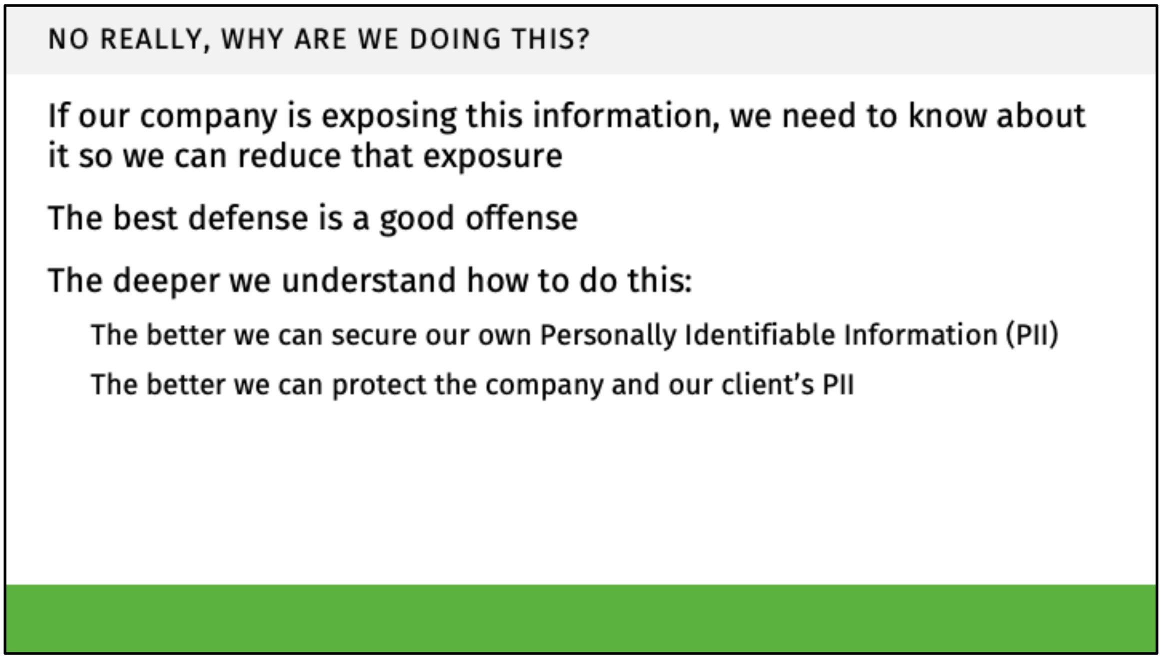
- Location of datacenter (if an employee posts while on the job)

The location of the data when it was made available

- Secret military base (Fitbit/Strava)

The author/creator of the document

- (domain username) perhaps



NO REALLY, WHY ARE WE DOING THIS?

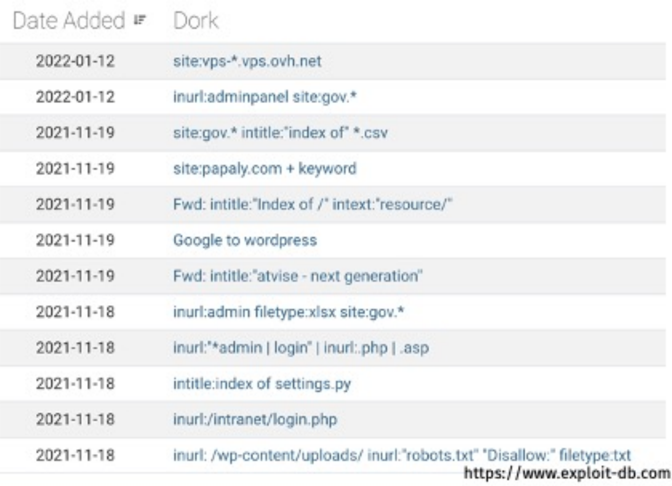
If our company is exposing this information, we need to know about it so we can reduce that exposure

The best defense is a good offense

The deeper we understand how to do this:

- The better we can secure our own Personally Identifiable Information (PII)
- The better we can protect the company and our client’s PII

BACK TO GOOGLE



<table>
<tr><th>Date Added</th><th>Dork</th></tr>
<tr><td>2022-01-12</td><td>site:vps-*.vps.ovh.net</td></tr>
<tr><td>2022-01-12</td><td>inurl:adminpanel site:gov.*</td></tr>
<tr><td>2021-11-19</td><td>site:gov.* intitle:"index of" *.csv</td></tr>
<tr><td>2021-11-19</td><td>site:papaly.com + keyword</td></tr>
<tr><td>2021-11-19</td><td>Fwd: intitle:"Index of /" intext:"resource/"</td></tr>
<tr><td>2021-11-19</td><td>Google to wordpress</td></tr>
<tr><td>2021-11-19</td><td>Fwd: intitle:"atvise - next generation"</td></tr>
<tr><td>2021-11-18</td><td>inurl:admin filetype:xlsx site:gov.*</td></tr>
<tr><td>2021-11-18</td><td>inurl:"*admin | login" | inurl:.php | .asp</td></tr>
<tr><td>2021-11-18</td><td>intitle:index of settings.py</td></tr>
<tr><td>2021-11-18</td><td>inurl:/intranet/login.php</td></tr>
<tr><td>2021-11-18</td><td>inurl: /wp-content/uploads/ inurl:"robots.txt" "Disallow:" filetype:txt</td></tr>
</table>

https://www.exploit-db.com/google-hacking-database

SEARCH OPERATORS

site: shows results for a given website

ext: shows results with just a specific extension

Intext: searches for a string in text

Inurl: searches within the URL

Intitle: searches within the HTML page title

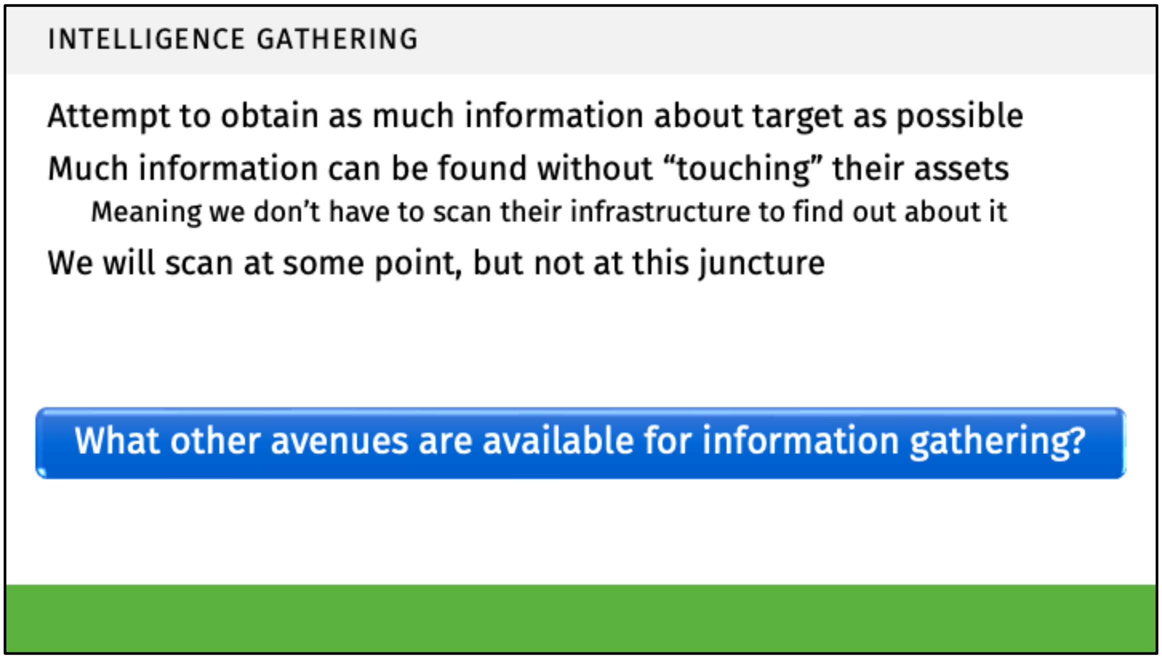
“ “ search for ONLY the string in those quotes

| or

Chain these together for maximum captchas benefit

- But yes, expect to get a captcha after a few of these searches

Select all images with a **store front**.
Click verify once there are none left.



INTELLIGENCE GATHERING

Attempt to obtain as much information about target as possible

Much information can be found without “touching” their assets

- Meaning we don’t have to scan their infrastructure to find out about it

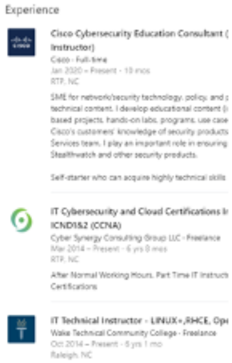
We will scan at some point, but not at this juncture

What other avenues are available for information gathering?

EMPLOYEE INFORMATION

How do we know who works at a company?

LinkedIn and social media, of course!



Experience

Cisco Cybersecurity Education Consultant (Instructor)
Cisco · Full-time
Jan 2020 – Present · 10 mos
RTP, NC

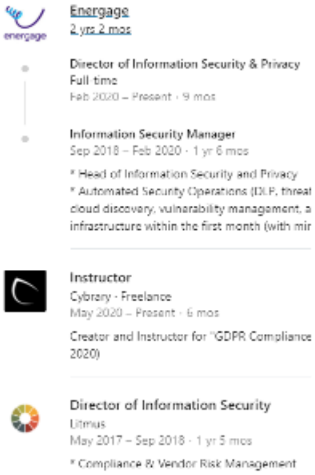
SME for network/security technology, policy, and s technical content. I develop educational content (i based projects, hands-on labs, programs, use case Cisco's customers' knowledge of security products Services team. I play an important role in ensuring Stealthwatch and other security products.

Self-starter who can acquire highly technical skills

IT Cybersecurity and Cloud Certifications In ICND1&2 (CCNA)
Cyber Synergy Consulting Group LLC · Freelance
Mar 2014 – Present · 6 yrs 0 mos
RTP, NC

After Normal Working Hours, Part Time IT Instructo Certifications

IT Technical Instructor - LINUX+,RHCE, Ope
Wake Technical Community College · Freelance
Oct 2014 – Present · 6 yrs 1 mo
Raleigh, NC



Energage
2 yrs 2 mos

Director of Information Security & Privacy
Full-time
Feb 2020 – Present · 9 mos

Information Security Manager
Sep 2018 – Feb 2020 · 1 yr 6 mos

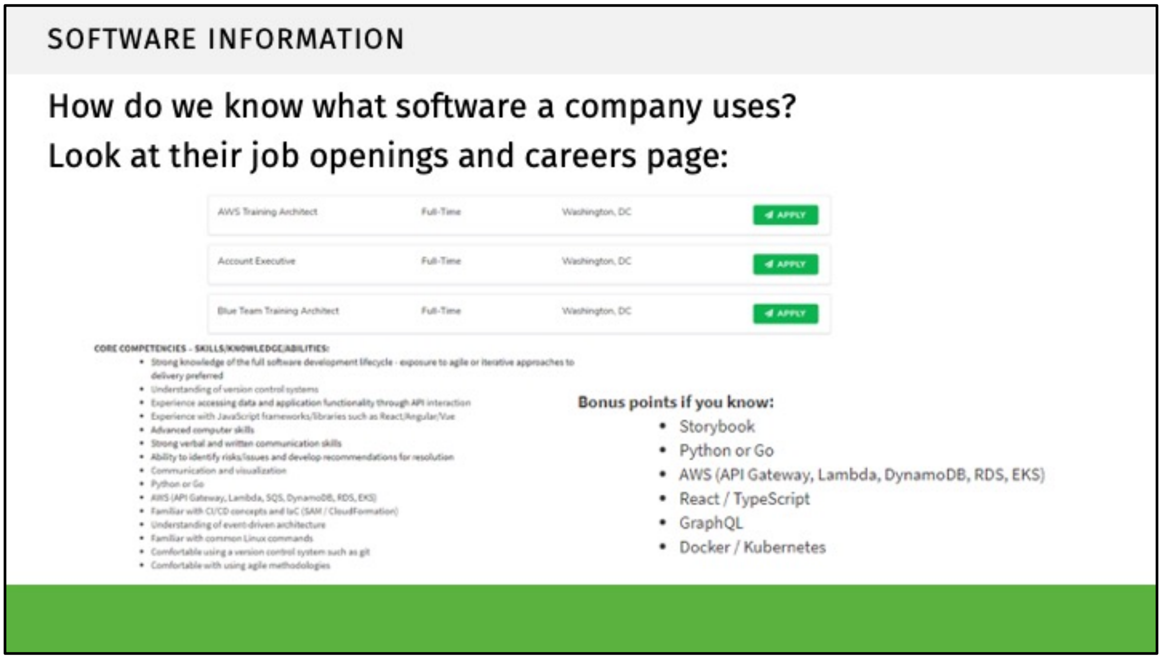
* Head of Information Security and Privacy
* Automated Security Operations (DLP, threat cloud discovery, vulnerability management, a infrastructure within the first month (with mir

Instructor
Cybrary · Freelance
May 2020 – Present · 6 mos

Creator and Instructor for "GDPR Compliance 2020)

Director of Information Security
Litmus
May 2017 – Sep 2018 · 1 yr 5 mos

* Compliance & Vendor Risk Management



SOFTWARE INFORMATION

How do we know what software a company uses?

Look at their job openings and careers page:

AWS Training Architect	Full-Time	Washington, DC	APPLY
Account Executive	Full-Time	Washington, DC	APPLY
Blue Team Training Architect	Full-Time	Washington, DC	APPLY

CORE COMPETENCIES – SKILLS/KNOWLEDGE/ABILITIES:

- Strong knowledge of the full software development lifecycle - exposure to agile or iterative approaches to delivery preferred
- Understanding of version control systems
- Experience accessing data and application functionality through API interaction
- Experience with JavaScript frameworks/libraries such as React/Angular/Vue
- Advanced computer skills
- Strong verbal and written communication skills
- Ability to identify risks/issues and develop recommendations for resolution
- Communication and visualization
- Python or Go
- AWS (API Gateway, Lambda, SQS, DynamoDB, RDS, EKS)
- Familiar with CI/CD concepts and IaC (SAM / CloudFormation)
- Understanding of event-driven architecture
- Familiar with common Linux commands
- Comfortable using a version control system such as git
- Comfortable with using agile methodologies

Bonus points if you know:

- Storybook
- Python or Go
- AWS (API Gateway, Lambda, DynamoDB, RDS, EKS)
- React / TypeScript
- GraphQL
- Docker / Kubernetes

NETWORK INFORMATION

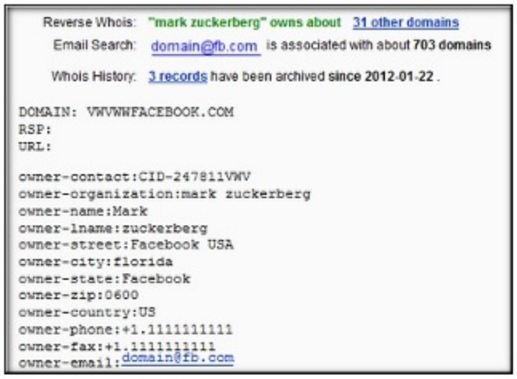
Where is their network hosted?

Who is their registrar?

Does their registrar have 2FA?

Check **whois** records ->

Check for domain AND IP

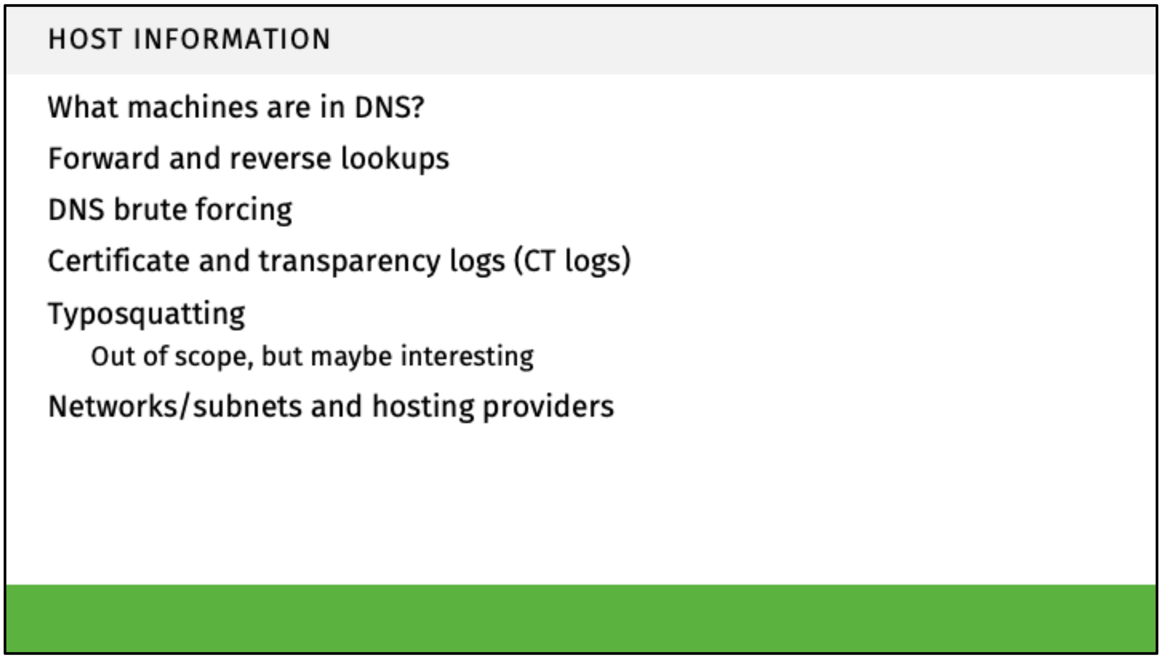


Reverse Whois: "mark zuckerberg" owns about 31 other domains

Email Search: domain@fb.com is associated with about **703 domains**

Whois History: 3 records have been archived **since 2012-01-22** .

```
DOMAIN: VWVWWFACEBOOK.COM
RSP:
URL:

owner-contact:CID-247811VWV
owner-organization:mark zuckerberg
owner-name:Mark
owner-lname:zuckerberg
owner-street:Facebook USA
owner-city:florida
owner-state:Facebook
owner-zip:0600
owner-country:US
owner-phone:+1.1111111111
owner-fax:+1.1111111111
owner-email:domain@fb.com
```



HOST INFORMATION

What machines are in DNS?

Forward and reverse lookups

DNS brute forcing

Certificate and transparency logs (CT logs)

Typosquatting

- Out of scope, but maybe interesting

Networks/subnets and hosting providers

LOCATION AND INDUSTRY

- Contact page on website + Google Maps is a good place to start
- We can go deeper:
  - What about location of their data center?
  - What about their data recovery/business continuity locations?
  - What about their company picnic location?
  - Do they sponsor golf tournaments?
  - How about their last Halloween party?
    - A little creepy maybe, but can be useful in understanding exposure

This Photo by Unknown Author is licensed under CC BY-SA

NETWORK SCANNING

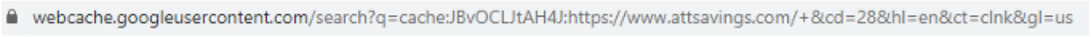
We’ll spend a whole day on this later

For now, know we can still find out a lot without touching their network

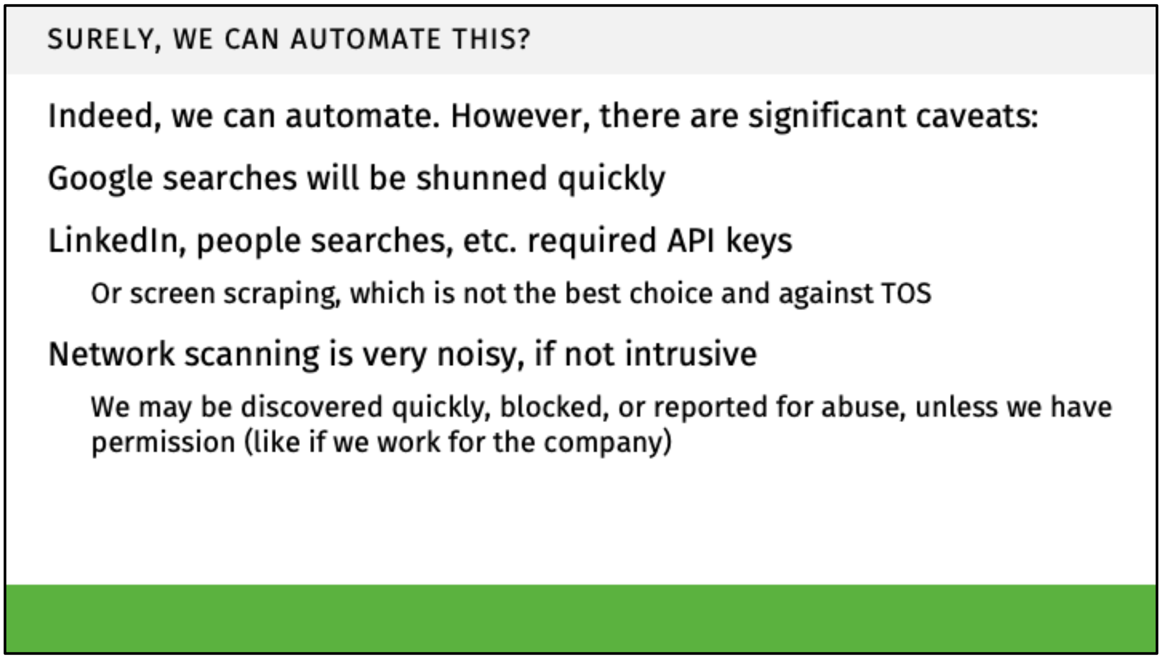
There are benefits to doing “offline” information gathering

Another thing Google is useful for (cached results):

- This allows us to visit websites via Google’s cached search results



webcache.googleusercontent.com/search?q=cache:JBvOCLJtAH4J:https://www.attsavings.com/+&cd=28&hl=en&ct=clnk&gl=us



SURELY, WE CAN AUTOMATE THIS?

Indeed, we can automate. However, there are significant caveats:

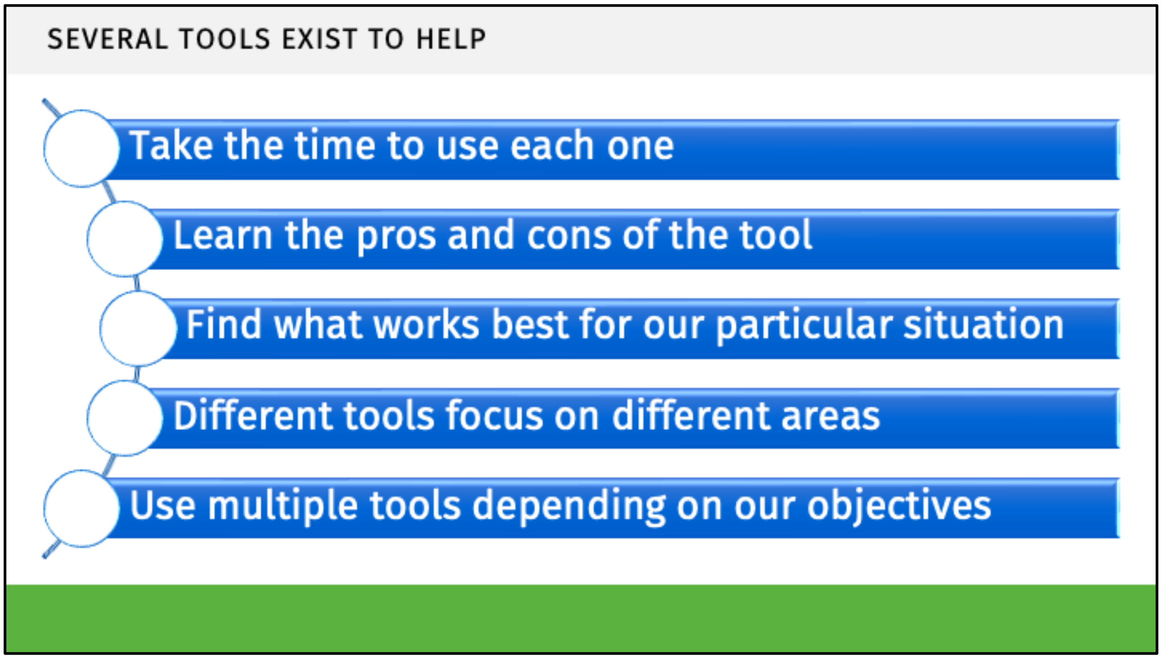
Google searches will be shunned quickly

LinkedIn, people searches, etc. required API keys

- Or screen scraping, which is not the best choice and against TOS

Network scanning is very noisy, if not intrusive

- We may be discovered quickly, blocked, or reported for abuse, unless we have permission (like if we work for the company)

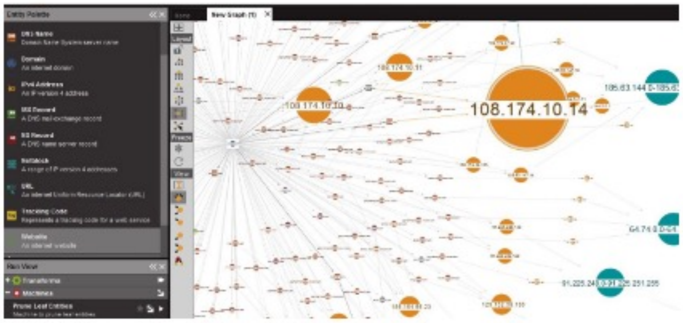


SEVERAL TOOLS EXIST TO HELP

- Take the time to use each one
- Learn the pros and cons of the tool
- Find what works best for our particular situation
- Different tools focus on different areas
- Use multiple tools depending on our objectives

MALTEGO

- Software used for open-source intelligence and forensics
- Mature software
- Free and paid versions
- Uses “transforms”



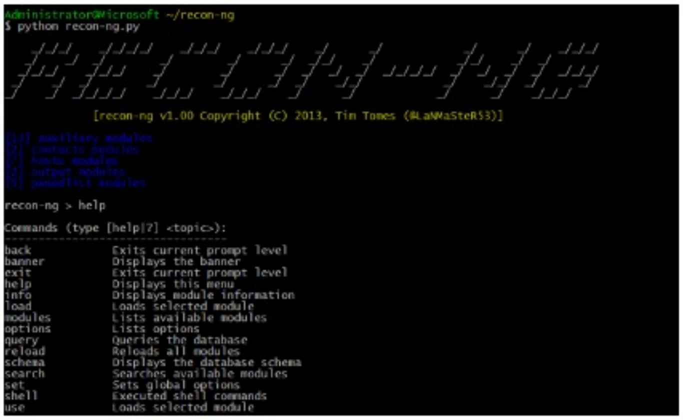
RECON-NG

Modular approach to OSINT

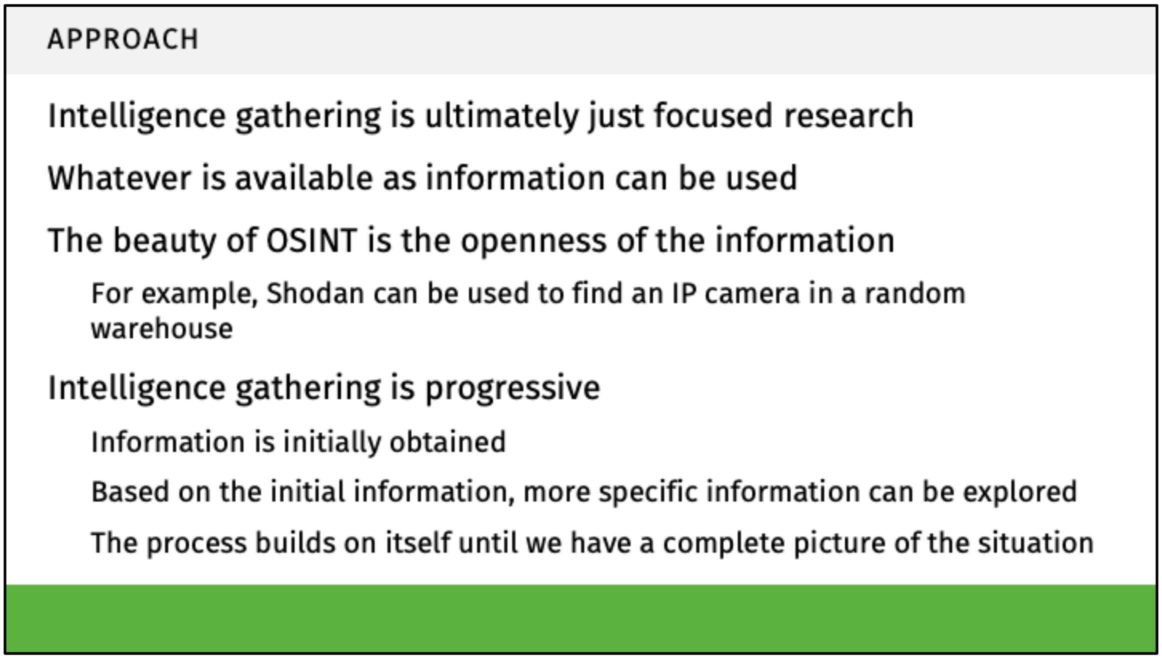
CLI

Web-based open-source reconnaissance

Also needs API keys for extended functionality

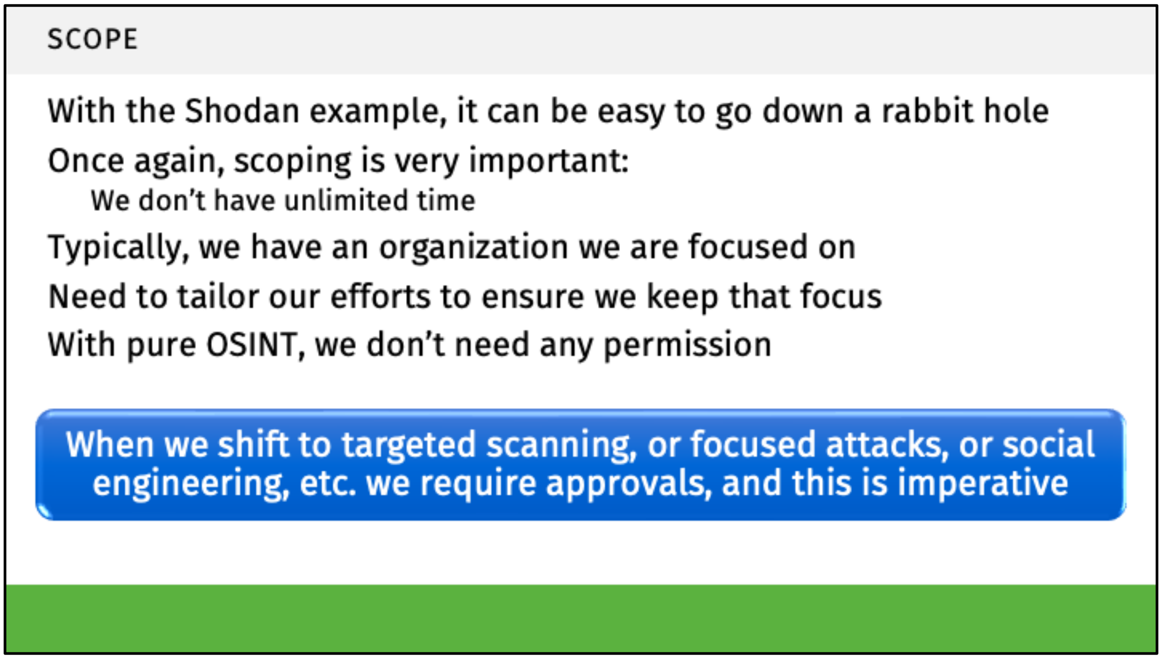


```
Administrator@Microsoft ~/recon-ng
$ python recon-ng.py

         [recon-ng v1.00 Copyright (C) 2013, Tim Tomes (@LaNMaSteR53)]

[11] auxiliary modules
[2] contacts modules
[7] hosts modules
[4] output modules
[5] pwnedlist modules

recon-ng > help

Commands (type [help|?] <topic>):
---------------------------------
back            Exits current prompt level
banner          Displays the banner
exit            Exits current prompt level
help            Displays this menu
info            Displays module information
load            Loads selected module
modules         Lists available modules
options         Lists options
query           Queries the database
reload          Reloads all modules
schema          Displays the database schema
search          Searches available modules
set             Sets global options
shell           Executed shell commands
use             Loads selected module
```



APPROACH

- Intelligence gathering is ultimately just focused research
- Whatever is available as information can be used
- The beauty of OSINT is the openness of the information
  - For example, Shodan can be used to find an IP camera in a random warehouse
- Intelligence gathering is progressive
  - Information is initially obtained
  - Based on the initial information, more specific information can be explored
  - The process builds on itself until we have a complete picture of the situation



SCOPE

With the Shodan example, it can be easy to go down a rabbit hole

Once again, scoping is very important:

- We don’t have unlimited time

Typically, we have an organization we are focused on

Need to tailor our efforts to ensure we keep that focus

With pure OSINT, we don’t need any permission

When we shift to targeted scanning, or focused attacks, or social engineering, etc. we require approvals, and this is imperative