

Practical OSINT

SecTor 2013

Shane MacDougall

~~Practical~~ OSINT

Offensive

DerbyCon 2013

Shane MacDougall

About Me

- .Been in InfoSec since 19eightyfreaking9holycrapthatiswaaaytoolong
- .Own Tactical Intelligence, a competitive intelligence firm (among other things)
- .Run the threat intelligence program for a F1000 company
- .Fair to middling social engineer
- .Writing a book on OSINT for NoStarch Press
- .Super awesome powerpoint skillz

WARNING

There is discussion of adult topics and (potentially) mild sexual content in this presentation. It is not gratuitous, and it relates directly to the presentation.

If there are kids in the room, you should remove them unless you don't mind answering uncomfortable questions later :)

If you are offended or affected by such content/talk, please take this opportunity to leave.

Outline

- .What is OSINT
- .Why OSINT
- .New/Underutilized Tools you should use
- .Offensive OSINT
- .Extortion-based OSINT
- .C2: Ceiling Cat Demo

What Is OSINT?

Open Source INTelligence

The hacker community uses a different definition of OSINT from what the .mil crowd does.

OSINT is defined by both the U.S. **Director of National Intelligence** and the U.S. **Department of Defense (DoD)**, as "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."

[Source: Wikipedia]

They don't include HUMINT,
TECHNINT...

HUMINT (Human Intelligence): Espionage, NGOs, Patrolling (Military police, patrols, etc.), strategic reconnaissance, as by Special Forces, Traveler debriefing (e.g., CIA Domestic Contact Service)

GEOINT

Geospatial Intelligence - gathered from satellite, aerial photography, mapping/terrain data

IMINT

Imagery Intelligence: gathered from satellite and aerial photography

OSINT

Open Source Intelligence can be further segmented by source type; Internet/General, Scientific/Technical and various HUMINT specialties (e.g. trade shows, association meetings, interviews, etc.)

They Don't Include

SIGINT

Signals Intelligence – intelligence-gathering by interception of [signals](#), whether between people ("communications intelligence"—COMINT) or from electronic signals not directly used in communication ("electronic intelligence"—ELINT), or a combination of the two. -

[Wikipedia]

COMINT - Communications Intelligence

ELINT - Electronic Intelligence: gathered from non-communications electronic emissions

TELINT - Telemetry Intelligence: the collection and analysis of telemetry data

TECHINT - Technical Intelligence.

FININT - Financial Intelligence - gathered from analysis of monetary transactions

How The Hacker Comm Defines It

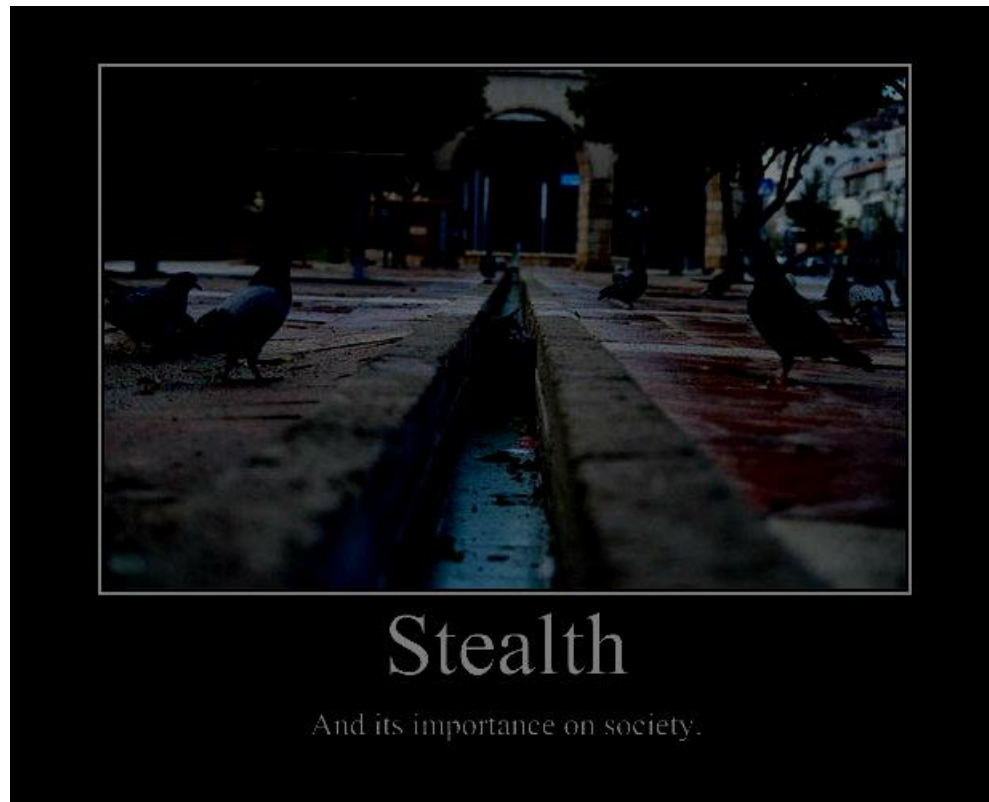
MOST OF THE ABOVE!

We use OSINT to grab the internet versions of the above from other intelligence/data aggregators (or we grab it ourselves via crawlers and other feeds)

We may bastardize some of the scenarios or actors from above, but basically we can find most of those data points (or those similar to) on the web.

Why OSINT?

Allows you to gather huge amounts of actionable intelligence without ever sending a packet to your target.



Why OSINT?

The purpose of OSINT is to optimize an attack, be it a network/application attack or a “kinetic” physical or social engineering attack. Attackers want to reduce the noise they make, and with enough profiling we can.

Password cracking/KBA optimization/Social Engineering:

- .pet name
- .year of birth
- .significant other/relatives names
- .sports team
- .religious phrase
- .place of birth
- .favorite food
- .where they were married
- .previous addresses
- .previous employers
- .list of schools
- .year account or mortgage was opened

We reduce the number of attempts and maximize our chances for success with every piece of information we gather.

Why OSINT?

OSINT can be devastatingly effective.

- BlackHat 2013

- Class had 45 minutes

- Broke into teams of 6 people

- Selected Haliburton as the target

Why OSINT?

- Teams discovered:
 - over 100 vulnerable running services
 - 1000's of employee SSN's, DoB's, email addresses, home addresses
 - credit card numbers
 - "black" site locations
- All without them sending a single packet to the target



Bad News

Changes to site TOS and efforts by companies to (gasp) make money over the past few years have seen a lot of great OSINT tools disappear.

Ongoing API changes and anti-crawling efforts on various sites mean some tools get broken and not fixed due to not being maintained.

Look at OSINT tools every year over the past 5 years. Breakage and deprecation abounds (frequently)!

IF YOU SEE SOMETHING SAY SOMETHING!!!

Report broken tools to the creators so they can fix them!



Good News

THE GOOD NEWS: There are always new tools coming along; some broken tools are getting fixed, and best of all? You can always roll your own tools!



Tools To Use

.Let's discuss some new (or less commonly used) OSINT tools that can make your life a lot easier.



Search Engines!!!



Yandex

- .Russia's largest search engine
- .Fourth largest in the world
- .Awesome search engine. Gives you great granularity with search operators.
- .Yet very few Hax0rs seem to use it.
- .Why?
- .It's Russian, and Mmmurica...
- .Use it!

Yandex

Like Bing and/or Google, it supports the inurl, site, link and lang operators, parentheses, and supports Boolean searches (like Bing it uses | instead of OR), wildcards, and (my personal favorite) implements the *+keyword* operator, which indicates the keyword **must** appear in the page.

Yandex: +

The search: kitten +mayor

means the word "mayor" must appear in the page, while kitten should appear (but it's not mandatory).



The screenshot shows the Yandex search engine interface. The search bar at the top contains the query "kitten +mayor". The left sidebar features navigation options: Web (selected), Images, Video, Mail, and Translation. The main content area displays two search results. The first result is a link to "Kitty Piercy. For Mayor. For Eugene." from kittypiercy.com, with contact information for Eugene Mayor Kitty Piercy. The second result is a link to "KITTEN-KILLING MAYOR ROCKS CANADA'S SELF-IMAGE - Myriad - Open..." from open.salon.com, dated 16 July 2013, with a snippet about a mayor apologizing for joking about killing cats.

Yandex

kitten +mayor

Web

Images

Video

Mail

Translation

[Kitty Piercy. For Mayor. For Eugene.](#)
kittypiercy.com
Contact Information for Eugene **Mayor Kitty Piercy**. City of Eugene 125 East 8th Avenue, 2nd Floor Eugene, OR 97401. Phone: (541) 682-5010 Email: kitty.piercy@ci.eugene.or.us.

[KITTEN-KILLING MAYOR ROCKS CANADA'S SELF-IMAGE - Myriad - Open...](#)
open.salon.com > blog...2013/07/16/kitten...mayor_rocks...
The **mayor** of Huntingdon, Que., Stephane Gendron, has been forced to apologize for joking about how he enthusiastically kills cats with his car — even newborns. Fearless **kitten**-killer (Photo of his...
16 July 2013

Yandex: ~ ~

~ ~ is the Yandex NOT, or exclusion operator.

new york ~ ~ city

will return web pages with new and york in them,
but not the word city.

[Note how unlike Bing and Google, there can be a
space after the operator.]

Yandex: ~

The ~ operator only returns sentences that exclude a keyword.

ice ~ cream

will return web pages containing any sentences containing "ice" but not "cream." Note that the pages can still contain "ice" and "cream," just not in the same sentence.

Yandex: & and &&

The & operator specifies that specified keywords must appear in the same sentence.

ying & yang

The && operator requires both words simply appear on the same webpage.

ying && yang

Yandex: */number*

The / operator is the Yandex equivalent of Google's AROUND and Bing's NEAR proximity search. It specifies how close together two keywords must be.

ice /3 cream

will only return results where "ice" and "cream" are within 3 words of each other (in any order, for example "ice" can come before "cream" or vice versa).

Yandex: */number*

By adding a + after the /, you can specify that the words be in the same order.

ice /+2 cream

will look for any webpage where the word ice precedes the word cream by no more than 2 positions. If cream appears before ice, it will not match the search.

Yandex: */number*

If you use a negative number, it will only return results where the second keyword appears X positions before the first.

cream /-1 ice

would return pages with "ice cream" but not "cream ice."

Yandex: */number*

Yandex lets you take it even further than that though; you can specify boundaries within which words can appear through the */(X Y)* operator. This allows you to specify the left and right offset of the second keyword from the first. For example,

ice */(-2 +3)* cream

will return results where cream precedes ice by 2 no more than positions, or appears no more than three words after ice.

Yandex: */number* and &&

Another interesting feature is that Yandex allows you to combine the */* operator with && to force the keywords to be within a certain number of sentences from each other.

"ice cream" && /2 sandwich

will only return web pages where "ice cream" is within 2 sentences from the word "sandwich." As with Google and Bing, the words can be in either direction of each other.

Yandex: !

The ! operator instructs Yandex to only return the exact form of the word provided. In other words, if you want to search on the word "cat" and only want results for that, and not variants such as cats, caterpillar, and so on, your search query would be !cat.

Yandex: !

The ! operator instructs Yandex to only return the exact form of the word provided. In other words, if you want to search on the word "cat" and only want results for that, and not variants such as cats, caterpillar, and so on, your search query would be !cat.

Yandex: <<

This operator allows you to tag on additional search terms without impacting on the rankings of the website in the results.

Yandex refers to this as a "non-ranking AND."

spare ribs sauce << Scotland

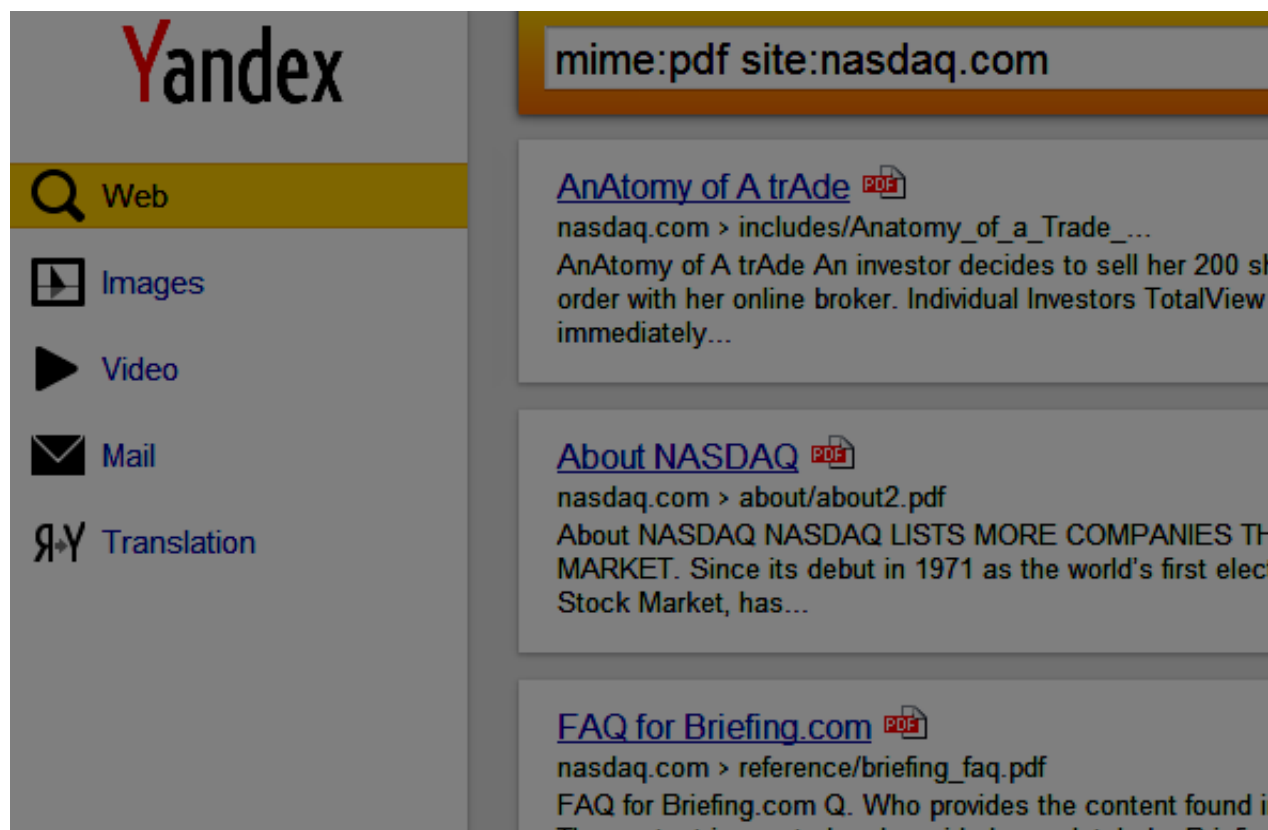
will not penalize a page for having the word Scotland associated with spare ribs (even though it probably should).

Yandex: `title:keyword(s)`

The `title` command is the Yandex equivalent of `intitle`, and returns only pages whose title contains the `keyword(s)`.

Yandex: mime:filetype

The Yandex equivalent of the filetype command, mime instructs the search engine to only return files that are of the type file_type.



The screenshot shows the Yandex search interface. The search bar contains the query "mime:pdf site:nasdaq.com". The left sidebar shows navigation options: Web (selected), Images, Video, Mail, and Translation. The main search results area displays three PDF files from nasdaq.com:

- [AnAtomy of A trAde](#) PDF
nasdaq.com > includes/Anatomy_of_a_Trade_...
AnAtomy of A trAde An investor decides to sell her 200 st order with her online broker. Individual Investors TotalView immediately...
- [About NASDAQ](#) PDF
nasdaq.com > about/about2.pdf
About NASDAQ NASDAQ LISTS MORE COMPANIES TH MARKET. Since its debut in 1971 as the world's first elec Stock Market, has...
- [FAQ for Briefing.com](#) PDF
nasdaq.com > reference/briefing_faq.pdf
FAQ for Briefing.com Q. Who provides the content found i

Yandex: date=date(s)

To set specific search windows, utilize the date command in one of the following methods:

date=201403* (results from March 2013)

date=>20130401 (results from > April 1, 2013)

date=20140320..20140704

(results from March 20, 2014 to July 4, 2014)

Yandex: domain:tld

Restricts the search to keywords within a certain top level domain (aka TLD, such as .com, .net, .ca, d so on). So the search

parrots domain:ca

will only return results on parrots from Canadian websites.

Yandex API

Very comprehensive list of API's for the various Yandex products:

api.yandex.com

Yandex: USE IT!

search.nerdydata.com

• "A Search Engine For Source Code"

• Unlike other search engines that index content, nerdydata indexes the HTML, Javascript, CSS, AND plaintext of over 140 million websites.

• "Give us any keyword or code snippet, and we'll return you a list of websites that contain it."

• If your target uses a specific Google Analytics account ID or javascript library, you can download the list of domains that contain it.

• Very easy to track unique backlinks

Google Hacking Obsolete?

• Instead of looking at the URLs or text within websites to determine whether they are vulnerable, we can now query the markup.

• `<meta name="generator" content="WordPress 3.5" />`

• `https://search.nerdydata.com/search/#!/searchTerm=<meta name="generator" content="WordPress 3.5" />/searchPage=1/sort=pop`

search.nerdydata.com

- Find pages with an upload form:

- Query: name="MAX_FILE_SIZE

- https://search.nerdydata.com/search/#!/searchTerm=name=%22MAX_FILE_SIZE%22/searchPage=1/sort=pop

search.nerdydata.com

- .Websites using the Invision Power Board Forum
- .Query: ipsBadge
- .<https://search.nerdydata.com/search/#!/searchTerm=ipsBadge/searchPage=1/sort=pop>

search.nerdydata.com

- Offers basic and premium (paid) services
- Premium subscribers get to download their results in large data sets, and get access to the "refining tool" which allows for more advanced searching.

Recon-ng

- Written by Tim Tomes

- www.recon-ng.com

- Open source reconnaissance framework

- Fantastic tool, lots of modules (demo if we have time)

- Works very much like Metasploit (same command structure)

- Constantly being revised

- "Do a git pull every time you run it." - Tim Tomes

Tapir

- By Jonathan Cran (@jcran)
- Web-based alternative to Maltego
- Run hosted or locally
- www.github.com/pentestify
- Work in progress but very close to release

Tapir Demo



OpenBmap.org

.Crowdsourced wireless mapping site: bluetooth, wi-fi, cellular antennas (CDMA and LTE site data acquisition is a bit wonky)

.As of last night contained data from 194 countries, 631 networks, 36353 location area codes, 435071 cells of which 134926 are trusted

.Wifi data from 66 countries, 654416 wifi access points

OpenBmap.org

- .Licensed under Open Database License

- .Cell data:

 - openbmap.org/latest/cellular/raw/input_raw.zip

- .Wifi/bluetooth data:

 - openbmap.org/latest/cellular/raw/input_raw.zip

OpenBmap.org Wifi File

```
<logfile manufacturer="LGE" model="Nexus 4" revision="4.3"  
swid="Radiobeacon" swver="00.7.01" exportver="00.7.03" >
```

```
<scan time="20130909111530" >
```

```
<gps time="20130909111530" lng="6.92670561"  
lat="50.91474492" alt="95.80000305175781"  
hdg="40.400001525878906" spe="5.25" accuracy="10.0" />
```

```
<wifiap bssid="24:65:11:dd:51:c6"  
md5essid="dfa8327f5bfa4c672a4f9b38e348a70" ssid="homer"  
capa="[WPA-PSK-TKIP][WPA2-PSK-CCMP][WPS][ESS]" ss="-81"  
ntiu="2472"/>
```

```
<wifiap bssid="00:1c:28:1a:f2:36"  
md5essid="974c3da01254d0f6d4b745fe4531faaf"  
ssid="NETCOLOGNE-4861" capa="[WPA-PSK-  
CCMP+TKIP][WPA2-PSK-CCMP+TKIP][WPS][ESS]" ss="-68"
```


Last Warning...

We are about to get into some sexual content.
Feel free to leave (no harm, no foul)

A Plea

The following areas of discussion are very small, niches. If you want to work in this area, PLEASE coordinate with me. Otherwise the pool will quickly become polluted (and useless)



Ethics...



Ethics...

Some of the stuff we are about to discuss could really ruin someone's life, marriage, or cause them to come after you with a loaded shotgun.

This research was done for equal parts
lulz and PoC.

Tread carefully and don't be a dick.

Trip Down Memory Lane

Back on September 4, 2006, a troll named Jason Fortuny posted a (very) graphic W4M ad on Seattle's Craigslist "casual encounters" board seeking "str8 brutal dom muscular male."



Fortuny Experiment

The ad was definitely a bit brutal in nature; the pic and the description described a very aggressive BDSM fantasy scenario.



Fortuny Experiment

- .178 men responded

- .145 sent photos

- .Responses included work and personal email addresses, names, IM screen names, telephone numbers

Fortuny Experiment

- Fortuny published all the responses (unedited) on Encyclopedia Dramatica
- People quickly started identifying the respondents



Fortuny Experiment

- .Microsoft worker got outed (emailing CAS ads from work != good idea)
- .Some got divorced
- .Fortuny got death threats
- .Some sued Fortuny (and won)

Hmmmm...



How Easy Would It Be To Automate Blackmail Using OSINT??



Already Being Done Commercially

– www.mugshots.com

– www.bustermugshots.com

– www.justmugshots.com

–

Use SEO to be top sites.

Pay large fee to remove entry

– “Maxwell Birnbaum”

– Google has buried these sites after a NYT article

Already Being Done Commercially

– www.ripoffreport.com

– www.complaintsboard.com

– www.yelp.com

–

Pay money and our "arbitrators" will decide whether to remove/bury the results...

Already Being Done Commercially

.Revenge porn sites

-submityourex.com

-myex.com

-many many many others

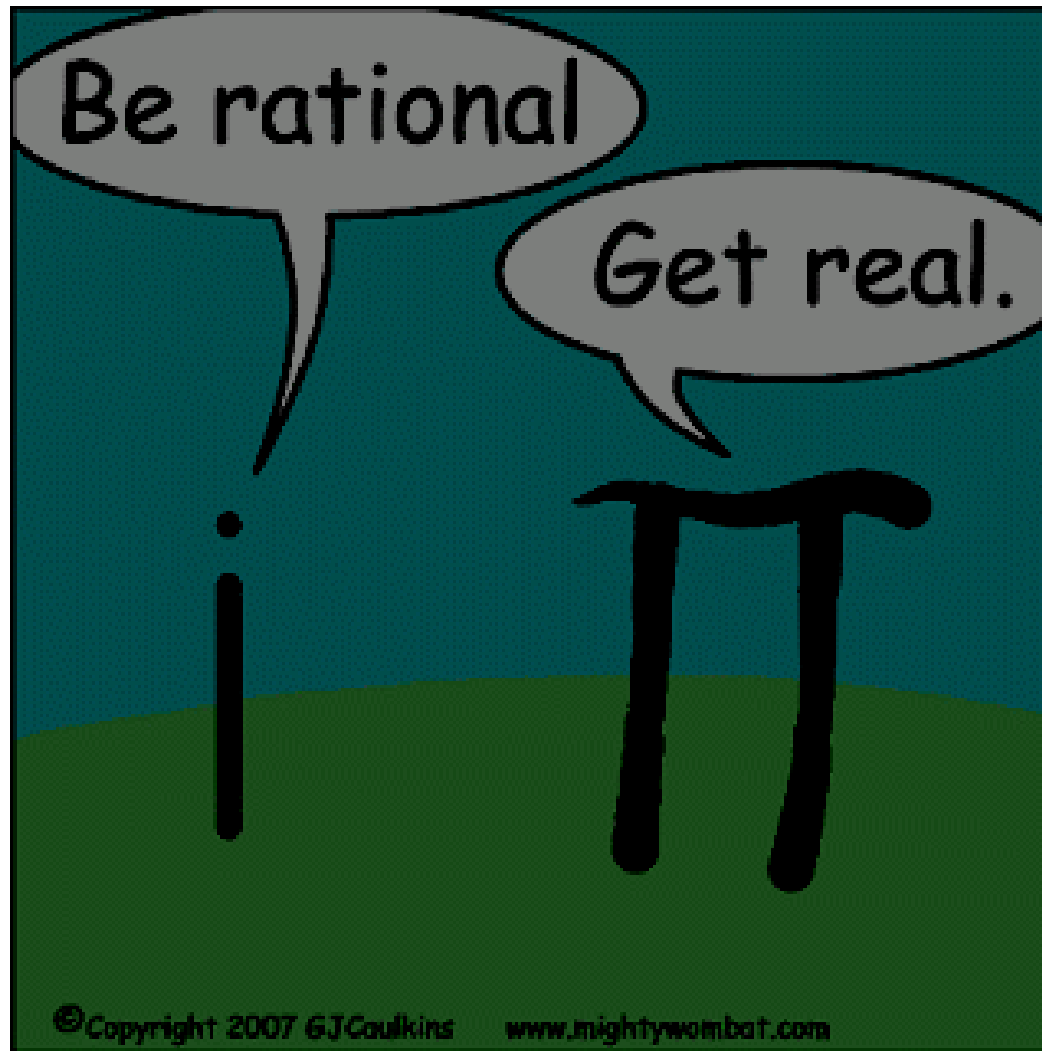
-

NJ and California have outlawed them, New York about to, but limited effectiveness. Step in the right direction.

DISCLAIMER

- .I really don't care what 2 (or more) consenting adults do in the privacy of their own homes.
- .The government (and media) really shouldn't care either
- .Sex is sex; as long as nobody is getting hurt (unwillingly) we should live and let live.

Sadly that's not how the world works



Fun With Fetlife

- Social network for fetishists
- "BDSM & Fetish Community for Kinksters by Kinsters"
- As of last night there were 2,385,192 members
- Potential goldmine for OSINT
- Lots of OpSec fail
- Some crossover with regular usernames (!!!)
- usernamecheck.com/namechk.com/etc
- LinkedIn trolling FTW!

Keep It In Your Pants!



Prostitution Sites

- Sites that allow "escorts" to sell their "wares"
- Very low bar to creating an account
- Grab a pretty picture, use a pre-paid credit card, create an account, and sit back and wait for johns to contact you. Harvest the info and profit...
- Providers don't have to vet themselves (other than click a "not LEO" box)
- The customers (often) do...

Provider Safety/Deadbeat Sites

Prostitution is a dangerous business. These ladies (and men) often share information on freaks, psychos, and "time wasters."

Gathering the information from these sites is pretty straight forward.

Unsurprisingly, no API for these sites. Need to scrape them.

Provider Safety/Deadbeat Sites

- .Remember these are crowdsourced sites, so YMMV
- .Reliability of the information "could" be suspect, although pretty unlikely.
- .Some require accounts/authentication/paid subscription
- .Those have a higher fidelity of data.
- .Some are wide open (and more suspect)

No Shows/Time Wasters



Sigh...
Stood Up Again

Damn, I Should've
joined Blacklist

**What's Your
Time worth?**

National BLACKLIST Deadbeat Registry
THE LARGEST BAD CLIENT REGISTRY SERVING THE ESCORT COMMUNITY
www.nationalblacklist.com

No Shows/Time Wasters

- .This information is usually freely available on most of the safety sites
- .The providers pay for the "dangerous" john info
- .Still, time wasters are people who are communicating with hookers...probably something they don't want their wives (or husbands or employers or constituents) to know about

Provider Safety/Deadbeat Sites

- <http://escortabuse.com/>

- Need an account to report, but not to search

- <http://blacklistedjohn.com/>

- Nice search feature, or you can crawl the entire database – limited content, hasn't been updated in 3 months

- <http://backpageblacklist.com/>

- Current, good search feature, paid to see incidents, time wasters are free

Provider Site in the EU

SAAFE.info

<http://www.saafe.info/main/index.php?topic=1615.0> (@saafe)

<http://www.saafe.info/main/index.php?board=3.0>
(warnings page – can harvest phone numbers, license plates, names(?), etc.

Provider Safety/Deadbeat Sites

<http://www.nationalblacklist.com/>

National BLACKLIST *Serving the Escorts Community*
DEADBEAT REGISTRY

- Home
- About BLACKLIST
- Our Cities
- Post A Report
- Subscribe to BLACKLIST
- Edit Your Account
- Browse Registry
- Search the Registry
- Automatic Alerts
- Online Payments
- Related News
- Articles & Tips
- Terms & Jargon
- FAQ
- Books & Resources
- Advertise on BLACKLIST

National BLACKLIST
DEADBEAT REGISTRY

Don't Get Fooled by Him Again!

For Escorts
World's Largest Bad Client Database and Escort Safety Tool!
67,793 Current Live Postings

BlackList Benefits...

Step 2 of 4 - Preferences
Choose your preferences for the automatic e-mail BlackList reports.

Select your region **West**

Yes, I want a report NOW

Yes, I want monthly reports

Yes, I want QuikALERTS!

Yes, I want the Newsletter

Grindr

- Grindr is a geosocial mobile app for gay/bisexual/bi-curious men
- Displays men looking to "hook up" that are in close proximity, and arranges the profiles/pictures in order of physical proximity to the user
- Users can choose to show their exact location (within feet!)
- Soooo...wandering around a target's neighborhood you may discover them in a closet (NTTAWWT)

Craigslist (duh!)

- Worked for Fortuny, still (surprisingly) works today
- Bad news is the dating site bots have taken over
- Still surprising the number of responses you get when you post as a female
- Like, really, really, really surprising
- And yes, they still email from work :((

Introducing C2: Ceiling Cat



C2: Ceiling Cat

- Twitter enumeration machine
- Written by myself and Akshit Agarwal
- Allows you to mine Twitter using regexes
- Can grab from the stream or via search
- The search module crawls Twitter (and breaks the ToS)
- API version coming (but much more limited)

C2: Ceiling Cat

- Geolocation (In Progress)
- From times of posts (like sleepingtime.org but without the bugs)
- Let's take a user Tim Smith (@bitvargen)
- <http://sleepingtime.org/>
- Now let's look at users Tim Smith (@clefnotes, @tim_smith20)
- Note the URL; multiple users

C2: Ceiling Cat - Geolocation

- .From weather reports (in process)
- .Reports of disasters
- .General Tweets (Location module)

C2: Ceiling Cat

- Comes preloaded with specific modules

- Format of modules:

 - # comment

 - "exact string"

 - (item 1 | item 2)

 - (optional)?

 - !(not)

 - (I | we | (my company) | (our company))

 - (use | uses) (windows | linux | macs | osx | debian | ubuntu)

C2: Ceiling Cat

.Has many uses:

-identify breaking news

-track specific topics

-identify characteristics about users

.End goal of the C2: Ceiling Cat project is to identify/profile all (or most) Twitter users and flag

C2: Ceiling Cat

- .Demo breaking news
- .Demo lifestyle

Next Steps

- .Create escort lookup modules for recon-ng, Maltego, and Tapir
- .Create "legit" API-based versions of C2
- .Distributed computing w/ C2
- .More languages for C2 (right now only English and Indian languages)
- .Connection to Twitter firehose for more comprehensive coverage

Contact Me

• shane@tacticalintelligence.org

• @tactical_intel

•

Thank you for your time!