

DEFINITIONS & ASSUMPTIONS



Dark Web

Refers specifically to a collection of websites that exist on an encrypted network and cannot be found by using traditional search engines or visited by using traditional browsers. Almost all sites on the so-called Dark Web hide their identity using the Tor encryption tool.

Deep Web

All of the web pages, or websites that have not been crawled by a search engine, are hidden behind paywalls or require a username and password to access. The opposite term is "surface web".

Dark Social

A term used by marketers to describe website referrals that are difficult to track. Dark social traffic doesn't seem to have a specific source, which creates a challenge for companies that are trying to monitor website referrals and social media activity.

Open Source Intelligence (OSINT)

Data collected from publicly available sources to be used in an intelligence context. In the intelligence community, the term "open" refers to overt, publicly available sources.

Social Media Intelligence (SOCMINT)

A subset of OSINT that gathers information exclusively from social media sites. It is typically analyzed from two layers – one, the original published content and two, the associated metadata.

Metadata

- Metadata is data about data. It's information that's used to describe the data that's contained in something like a web page, document, or file. A simple example of metadata for a document is author, file size, and the date created.
- Metadata comes in several types and is used for a variety of broad purposes that can be roughly categorized as business, technical, or operational.
 - **Descriptive** metadata properties include title, subject, genre, author and creation date, for example.
 - **Rights** metadata might include copyright status, rights holder or license terms.
 - **Technical** metadata properties include file types, size, creation date and time, and type of compression. Technical metadata is often used for digital object management and interoperability.
 - **Preservation** metadata is used in navigation. Example preservation metadata properties include an item's place in a hierarchy or sequence.
 - **Markup Languages** include metadata used for navigation and interoperability. Properties might include heading, name, date, list, and paragraph.
- It is used everywhere, by every industry, in multiple ways
- Metadata availability on social media depends on each platform (e.g. hashtags on Twitter are a public form of metadata). Metadata referring to the time/location of a photo or video is usually stripped upon upload.

ASSUMPTIONS



Investigations are being performed with public access to social data (SOCMINT) as opposed to private access (i.e. special access granted by a warrant).



It is understood that what your organization can/cannot do as a government organization is very different from what is theoretically possible.



It is understood that your organization should create a stand-alone protocol approved by proper legal counsel before using any of the tools in an official investigative capacity.



It is understood that the state of social media is constantly moving - what is/isn't possible changes on a daily basis. Any provided information may soon be out of date and need to be regularly updated.

POTENTIAL USES OF PUBLIC SOCIAL MEDIA DATA



Situational Awareness

Developing an overview of an unknown situation or event to focus future analysis



Ongoing Monitoring

Daily tracking of account activity and relevant keywords within a topic area



Social Network Analysis

Acquire an understanding of the social dynamics for large-scale data to tell you who and how users are relating



Priority Information Requests

Ask a pointed question to support an important decision



Investigations

Acquire additional information on a person or organization



Engagement

Analyze information around a communications goal or objective to provide engagement and content suggestions



Performance

Assess performance effectiveness and progress based on defined metrics measured against targets

 = focus of this session



2.5 quintillion (18 zeros) new data bytes produced daily

90% of the world's data has been produced in the last two years alone

Twitter still tends to be the most useful public data source for SOCMINT



**TIPS FOR
INVESTIGATIONS**



Basic operational security steps during an investigation

- Encrypt your connection
- Ideally use a browser dedicated to investigations
- Keep your operating system and browsers updated
- Ensure that you are using incognito or stealth mode, especially if you are not using the TOR browser
- Use anonymous search engines
- Use a dedicated password keeper
- Ensure your passwords are complex
- Run a virus and spyware scanner frequently
- Scan URLs you are unsure of
- Track your steps during an investigation





1 - SECURITY



PASSWORD SAFEKEEPING

1Password.com

Alternatives:

- Zoho Vault
- Dashlane
- LastPass
- LogMeOnce
- Password Boss

Purpose: Provides a place for users to store various passwords, software licenses, and other sensitive information in a virtual vault that is locked with a PBKDF2-guarded master password.



HOW SECURE IS MY PASSWORD?



It would take a computer about

42 MINUTES

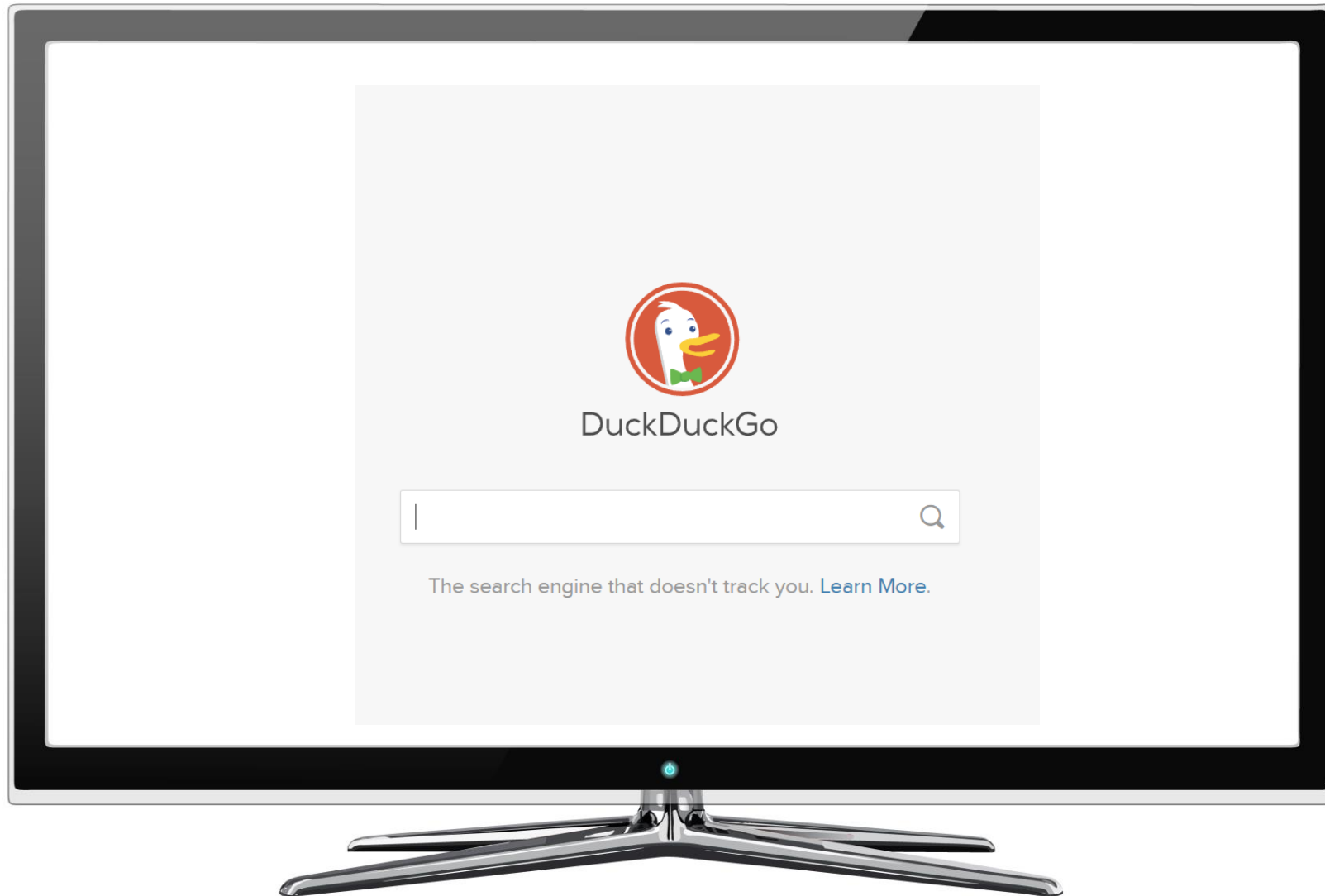
to crack your password

PASSWORD TESTING

howsecureismypassword.net

Purpose: A great way to see how easy it would be for someone to crack your password.



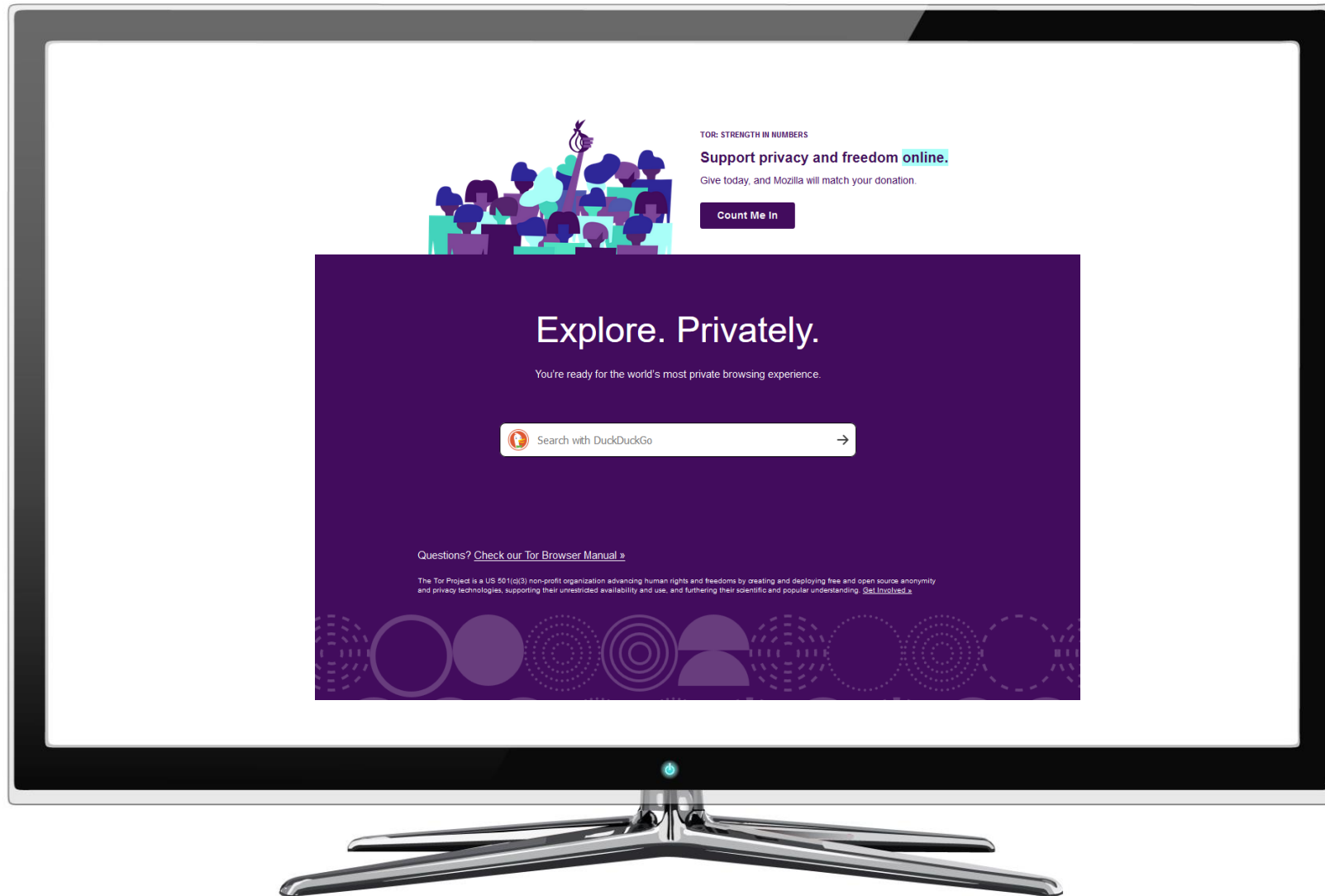


QUICK ANONYMOUS SEARCHING

duckduck.go

Purpose: Searching quickly
without being tracked



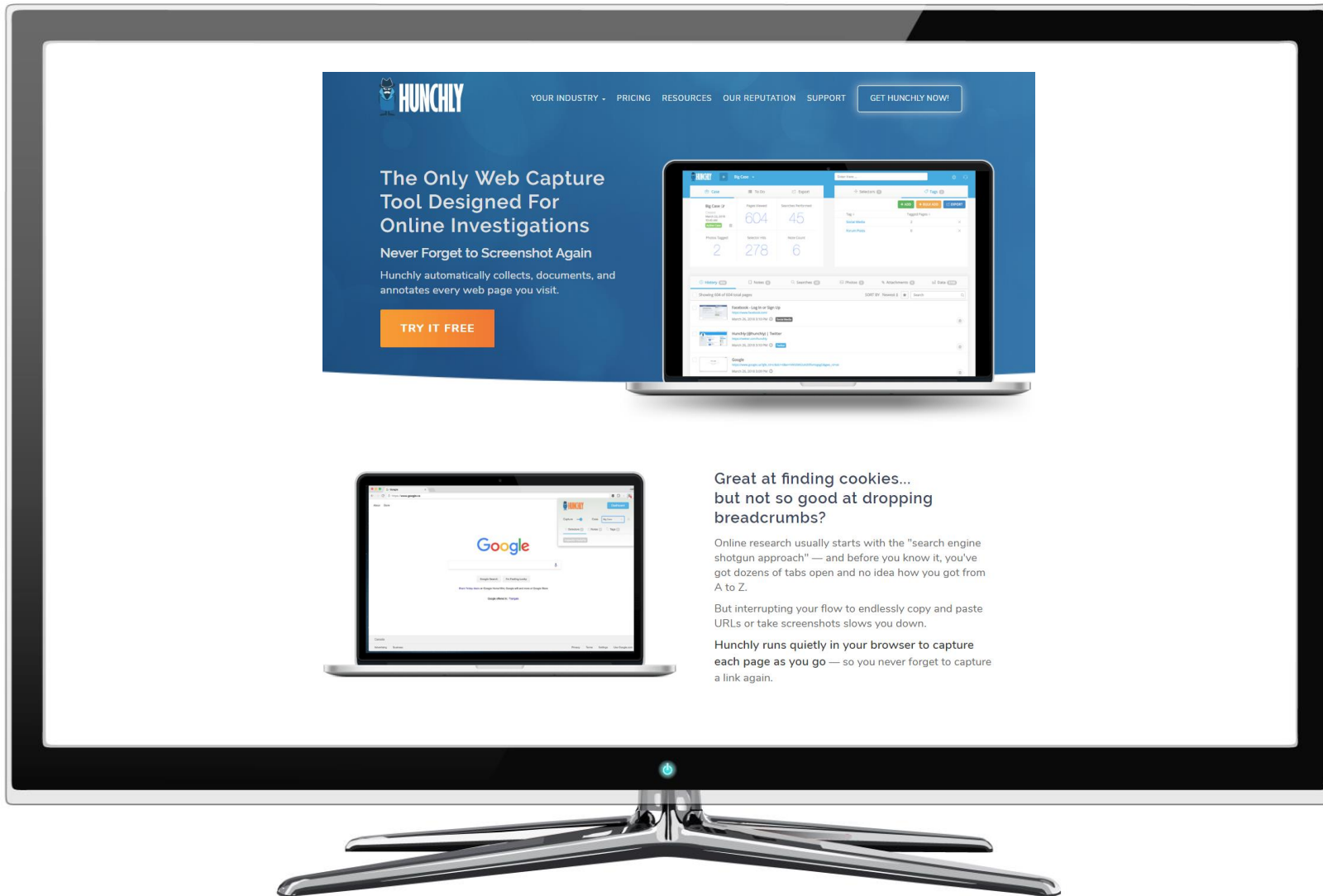


PRIVATE DARK WEB ACCESS

TOR Browser

Purpose: Useful when anonymity is of paramount importance. In the case of RoP it can be used to access a TOR compliant version of Facebook (facebookcorewwi.onion)





CAPTURING YOUR INVESTIGATION

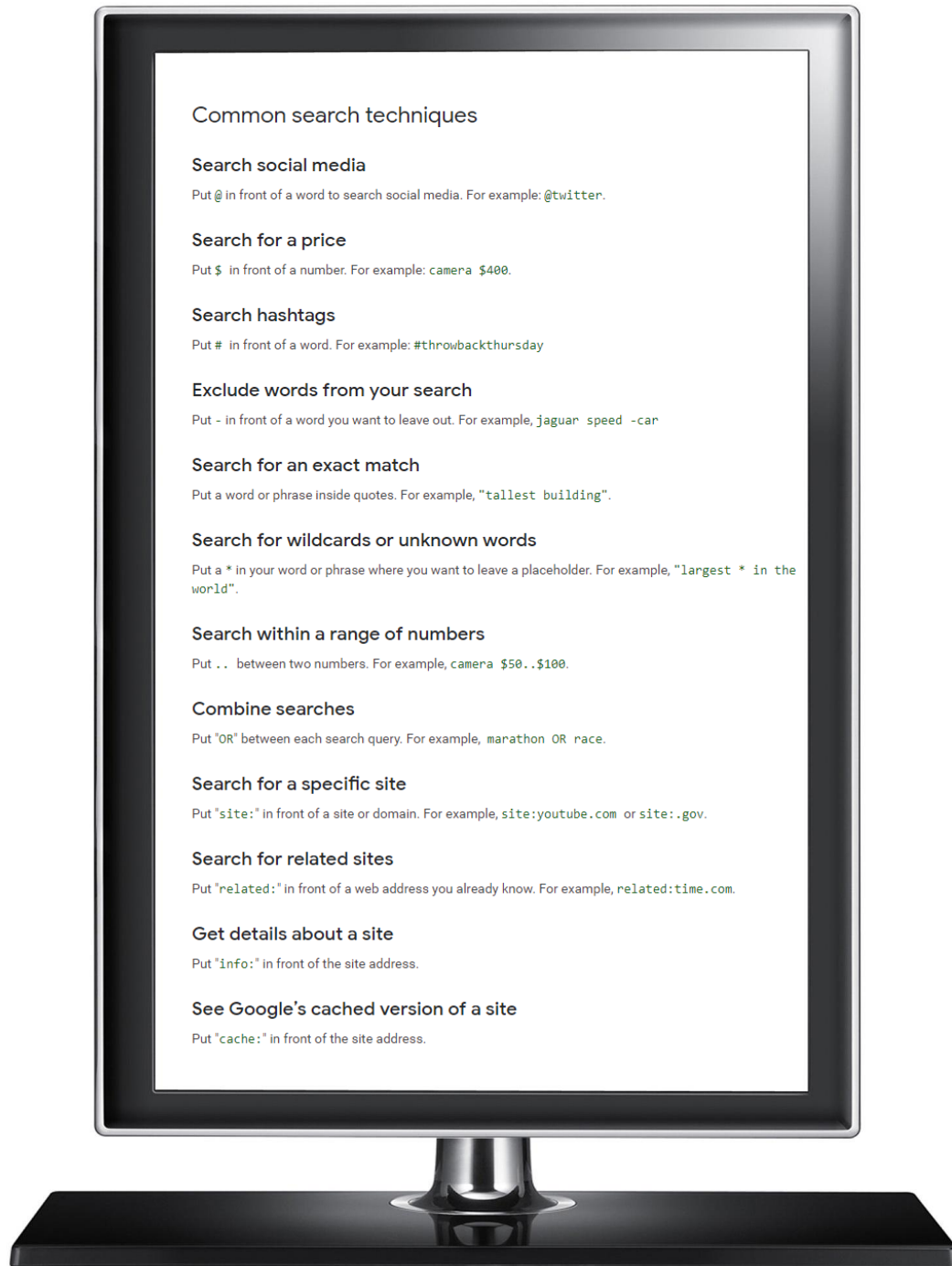
Hunchly.com
(Chrome Extension)

Purpose: Runs in the background during your investigation to capture each page as you go.





2 — ADVANCED GOOGLING



Common search techniques

Search social media

Put @ in front of a word to search social media. For example: @twitter.

Search for a price

Put \$ in front of a number. For example: camera \$400.

Search hashtags

Put # in front of a word. For example: #throwbackthursday

Exclude words from your search

Put - in front of a word you want to leave out. For example, jaguar speed -car

Search for an exact match

Put a word or phrase inside quotes. For example, "tallest building".

Search for wildcards or unknown words

Put a * in your word or phrase where you want to leave a placeholder. For example, "largest * in the world".

Search within a range of numbers

Put .. between two numbers. For example, camera \$50..\$100.

Combine searches

Put "OR" between each search query. For example, marathon OR race.

Search for a specific site

Put "site:" in front of a site or domain. For example, site:youtube.com or site:.gov.

Search for related sites

Put "related:" in front of a web address you already know. For example, related:time.com.

Get details about a site

Put "info:" in front of the site address.

See Google's cached version of a site

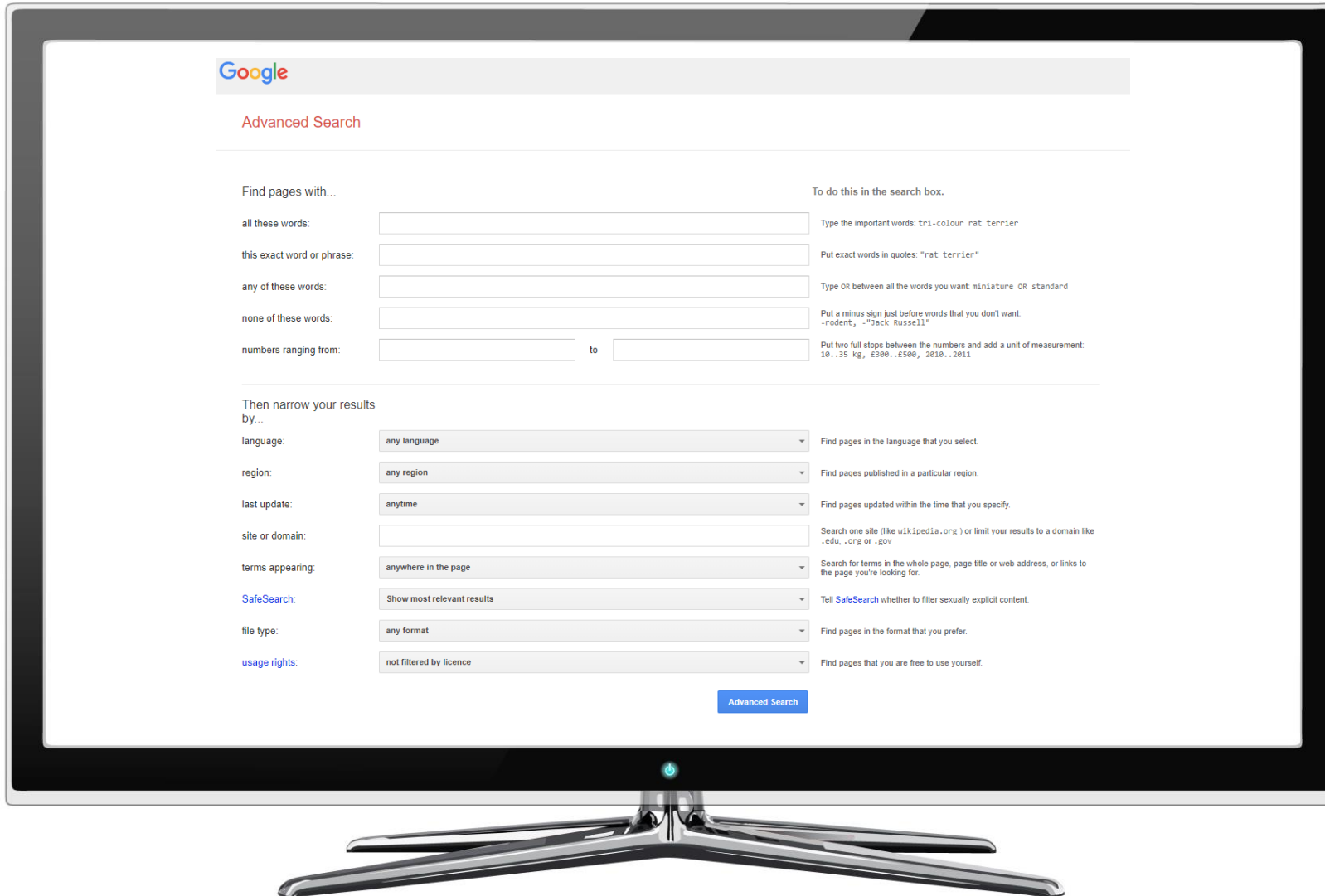
Put "cache:" in front of the site address.

BASIC SEARCH OPERATORS

goo.gl/4XfNU8

Purpose: Become faster at finding what you are looking for.





GOOGLE ADVANCED SEARCH

[google.ca/advanced
_search](https://google.ca/advanced_search)

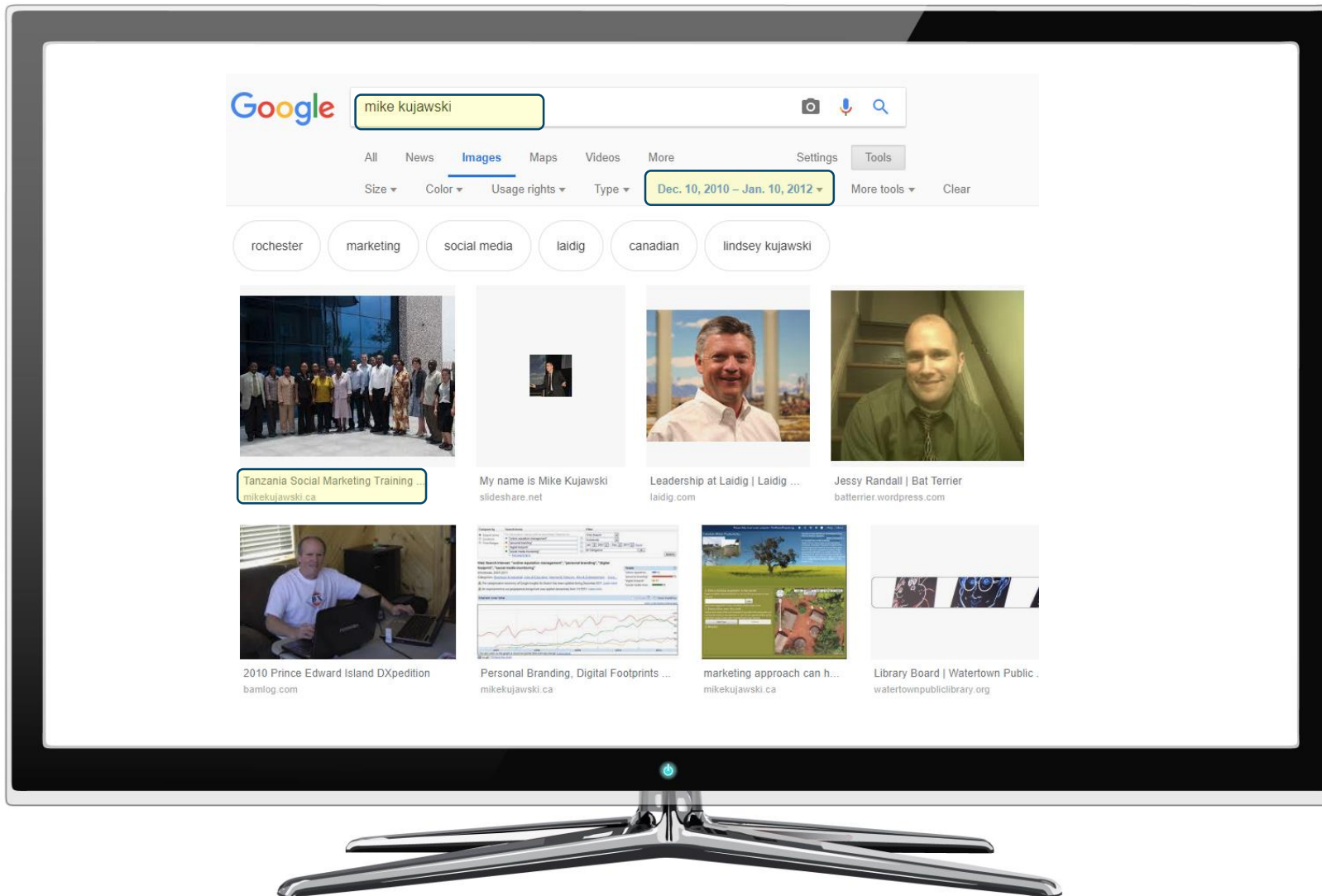
Purpose: For people that don't have time to memorize search operators



SEARCH BY SPECIFIC DATE

google → search
tools → date → custom
range

Purpose: To effectively search through historical Google data. Great for finding content that no longer appears in the top results.





**3 — VERIFYING
LEGITIMACY**

TinEye Technology Products Login

Reverse Image Search

Search by image and find where that image appears online

Upload or enter image URL

How to use TinEye



TinEye

Upload or enter Image URL

49 results

Searched over 31.7 billion images in 0.9 seconds.

for: <https://oceanicexplorer.files.wordpress.com/2013/01/dange...>

Oldest Filter by domain/collection < 1 of 5 >

img70229.pixa.us

Filename: 871b00dac3e6018e1aa015587e9d88c7.th.jpg

Found on: [images/5210959/](#)
Page crawled on Nov 07, 2009

Found on: [images/5217889/](#)
Page crawled on Nov 07, 2009

Compare Match

yachtingmonthly.com

Filename: DSC00205.JPG_e_fc7518c262fc1105be8a1c2ef5d90ab0.JPG

Found on: [news/500479/yachtsman-rescues-seven-d...](#)
Page crawled on Sep 25, 2010

Found on: [news/513517/petition-started-to-halt...](#)
Page crawled on Jan 20, 2011

Compare Match

dangerous-animals-pets.blogspot.ca

Filename: dangerous%2Bgreat%2Bwhite%2BShark-Attack%2Banimal%2Battac...

Found on: [2011/10/great-white-sharks.html](#)
Page crawled on Nov 30, 2013

Compare Match



RESEARCHING THE HISTORY OF A PHOTO

tineye.com

Purpose: To find all the places that a particular image appears online including the original source (even if it was modified)



VIEWING PHOTO METADATA

exifdata.com

Purpose: Viewing image
metadata on the go



SUMMARY
DETAILED
UPLOAD

System

File Name	1545163355.6_.jpg
File Size	223 kB
File Modify Date	2018:12:18 15:05:21-05:00
File Permissions	rw-r--r--

File

File Type	JPEG
MIME Type	image/jpeg
Image Width	1200
Image Height	900
Encoding Process	Progressive DCT, Huffman coding
Bits Per Sample	8
Color Components	3
Y Cb Cr Sub Sampling	YCbCr4:2:0 (2 2)

JFIF

JFIF Version	1.01
Resolution Unit	None
X Resolution	1
Y Resolution	1

Composite

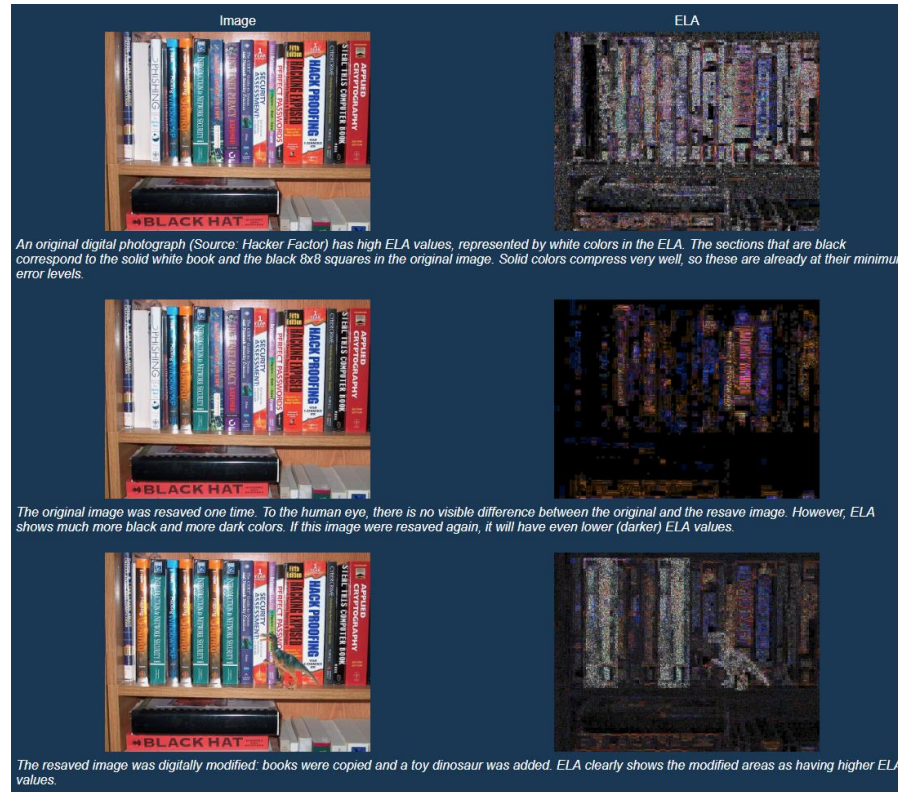
Image Size	1200x900
------------	----------



DETECTING FAKE PHOTOS

fotoforensics.com

Purpose: To determine if there were any modifications to an image

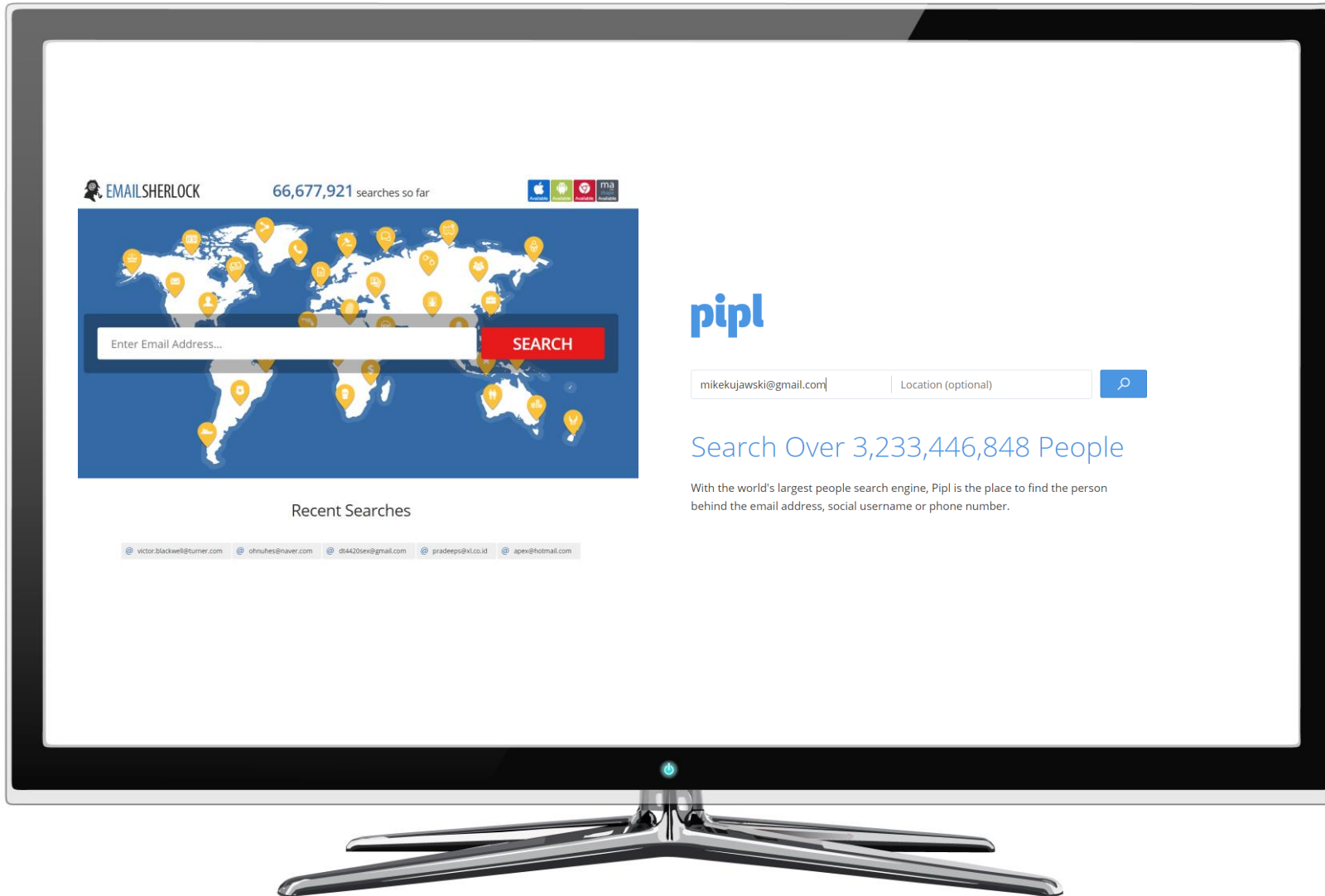


REVERSE EMAIL SEARCH

emailsherlock.com

pipl.com

Purpose: To help determine the identity of the email owner through scanning other places the email has been used online. Also great for discovering fraudulent emails.





Analyze suspicious files and URLs to detect types of malware, automatically share them with the security community.

File **URL** Search

By submitting your file to VirusTotal you are asking VirusTotal to share your submission with the security community and agree to our [Terms of Service](#) and [Privacy Policy](#). [Learn more](#).



No engines detected this URL

URL <https://www.canva.com/>
Host www.canva.com
Last analysis 2018-10-16 10:31:15 UTC
Community score +2

0 / 67

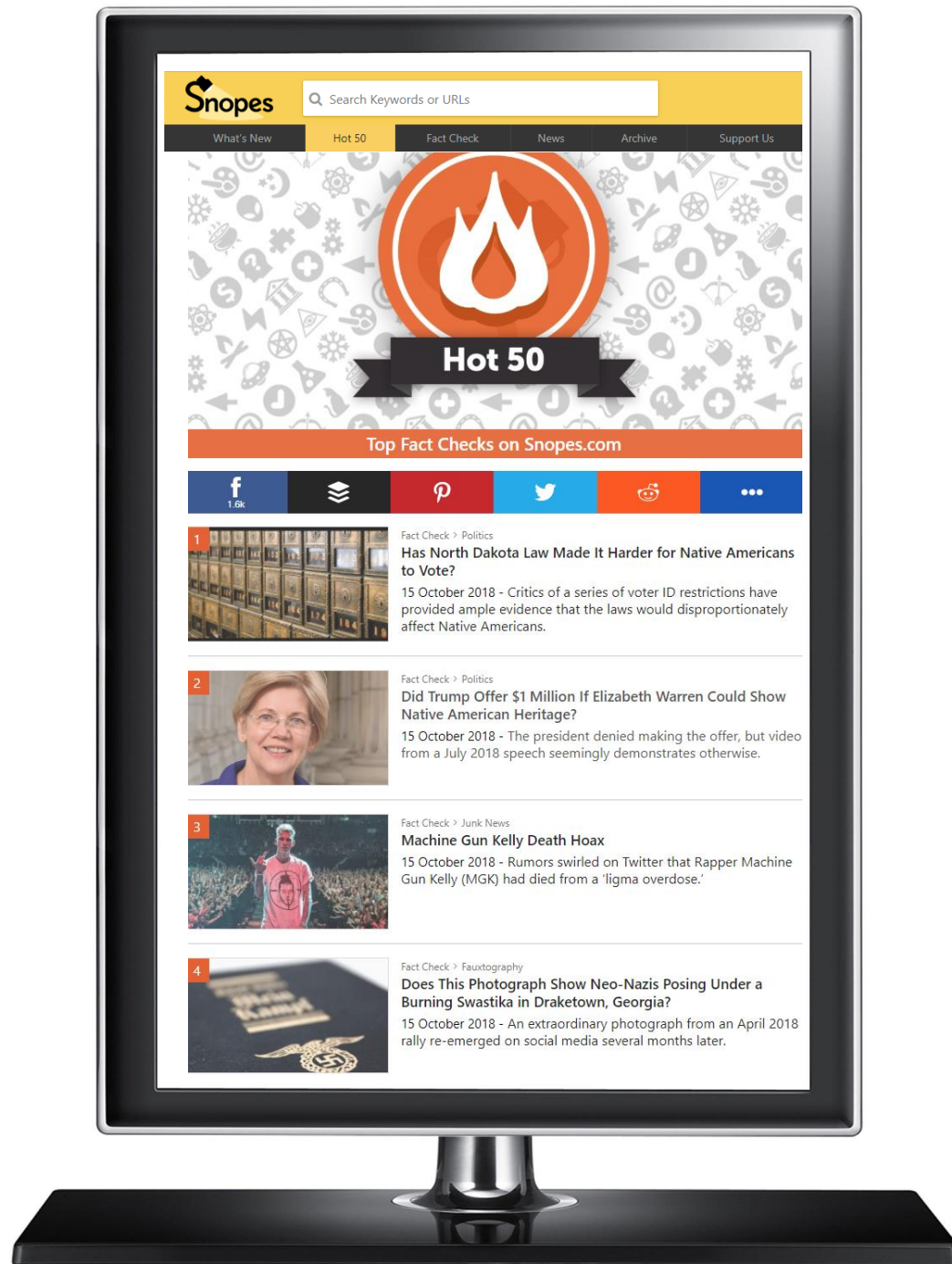
Detection	Details	Community
ADMINUSLabs	Clean	AegisLab WebGuard Clean
AlienVault	Clean	Antiy-AVL Clean
Avira	Clean	Baidu-International Clean
BitDefender	Clean	Blueliv Clean
C-SIRT	Clean	Certy Clean
CLEAN MX	Clean	Comodo Site Inspector Clean
CyberCrime	Clean	CyRadar Clean
desenmascara.me	Clean	DNSB Clean
Dr.Web	Clean	Emsisoft Clean
ESET	Clean	Forcepoint ThreatSeeker Clean
Fortinet	Clean	FraudScore Clean

CHECKING SUSPICIOUS LINKS

[virustotal.com](https://www.virustotal.com)

Purpose: Analyze suspicious files and URLs to detect types of malware, automatically share them with the security community.





BREAKING NEWS VERIFICATION

snopes.com

Purpose: To help prevent the spread of disinformation. Always use this before sharing a breaking story that is hard to believe.





4 - SOCIAL SEARCH

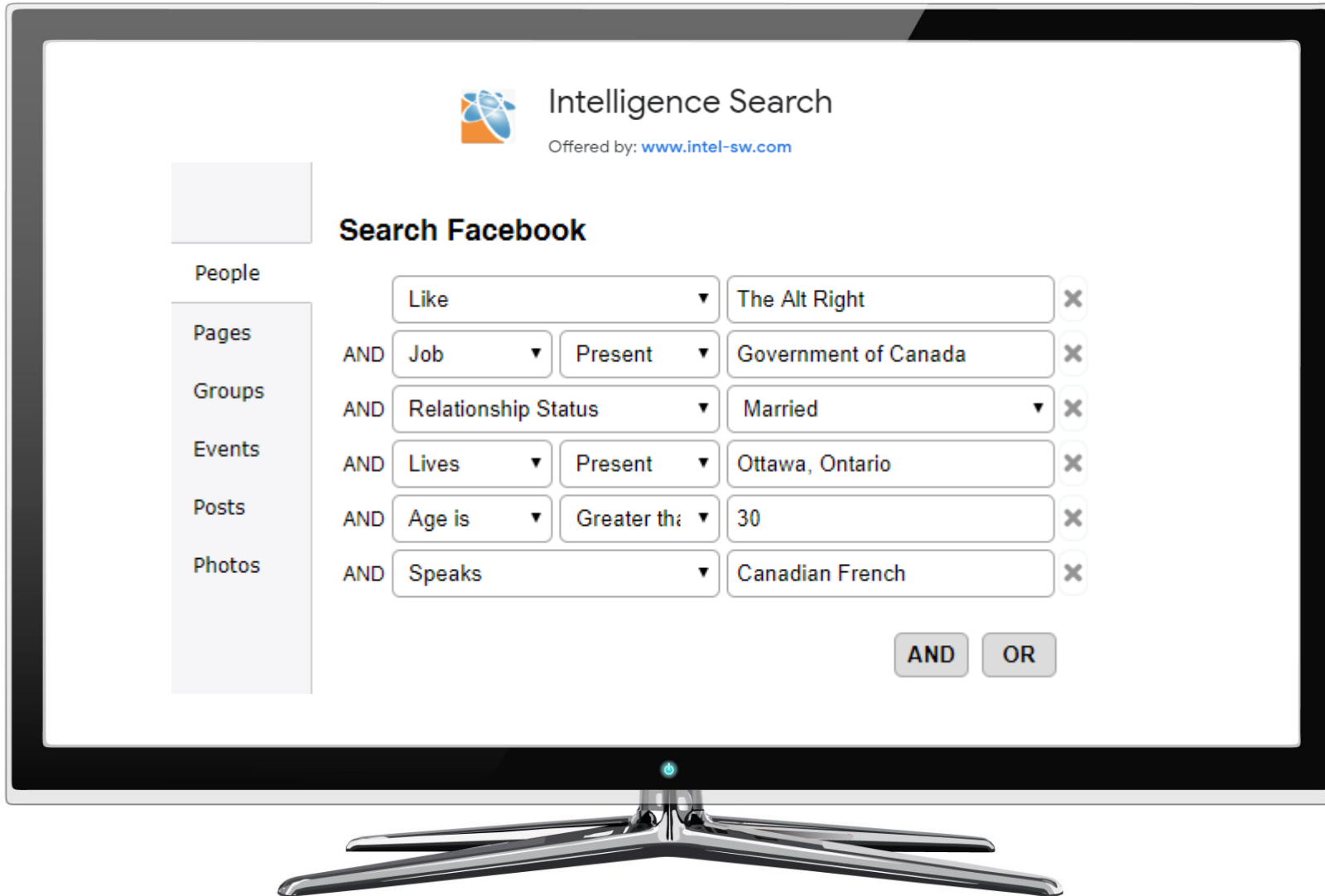
DIGITAL FOOTPRINT SEARCH 1

inteltechniques.com

***Use findmyfbid.com to find an ID number**

Purpose: To search through FB, TW, LI and find people by name, job, location, age, gender, their friends, the groups they are members of etc. Also great for managing your own digital footprint.



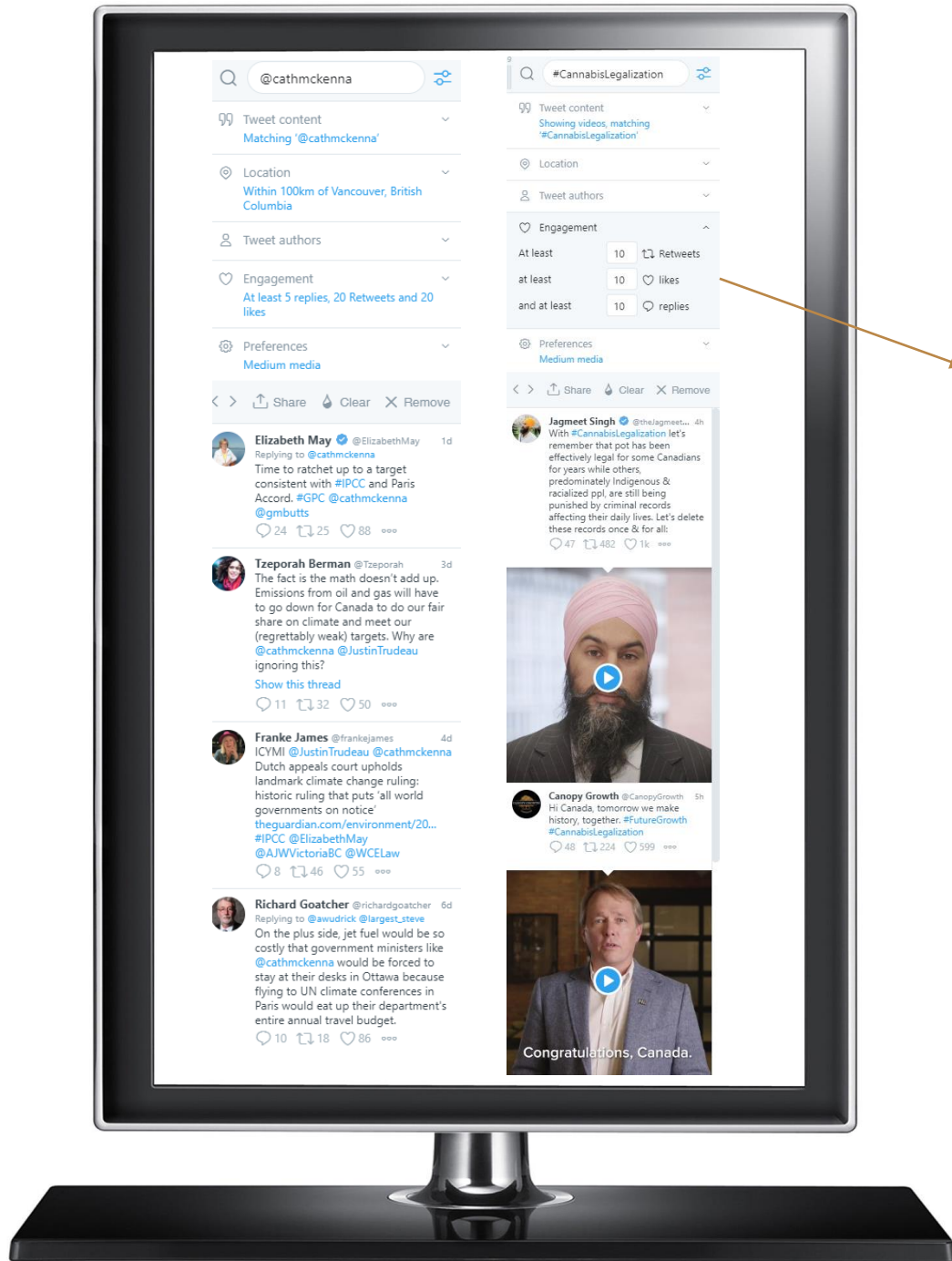


DIGITAL FOOTPRINT SEARCH 2

goo.gl/x4RPbK (Chrome Plugin)

Purpose: To search through FB, TW, LI and find people by name, job, location, age, gender, their friends, the groups they are members of etc. Also great for managing your own digital footprint.





Engagement


At least Retweets

at least likes

and at least replies

Location

Tweets geo-tagged near



Radius

Tweet content

Showing

Matching

Excluding

From

To

Written in

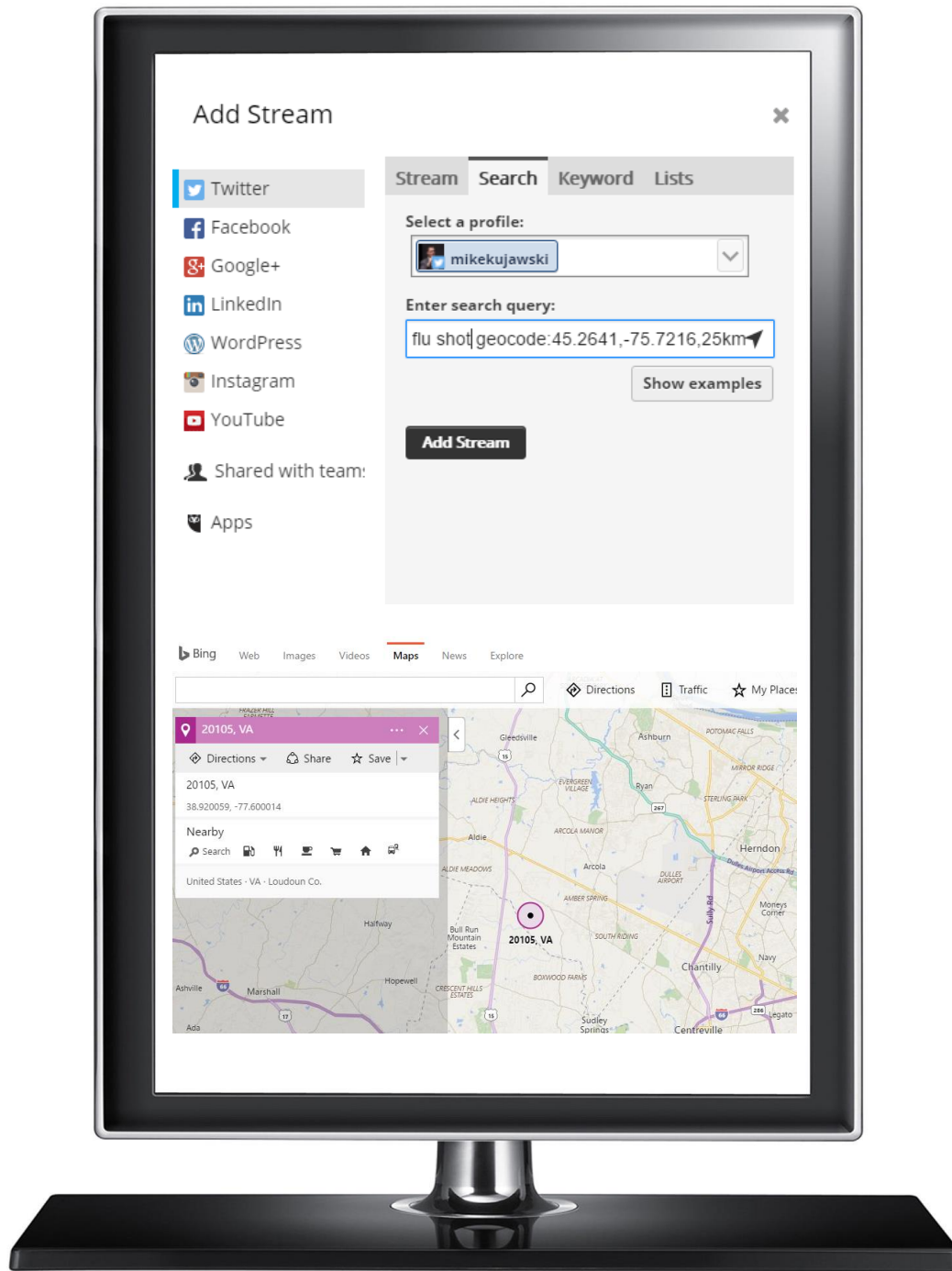
Retweets

REAL-TIME CONTENT FILTERING

tweetdeck.com

Purpose: To quickly filter content based on media type, engagement level, geo-location, etc. in real-time





GEO-LOCATING TWEETS USING HOOTSUITE

hootsuite.com

Purpose: To locate tweets in a specific area

Geo Location Operator

[search term] geocode:[latitude,longitude],[radius]km





TWITTER ACCOUNT ANALYSIS

twitonomy.com

Purpose: To verify accounts and better understand how they interact and who they influence.





GEOGRAPHY BASED TREND DISCOVERY

trendsmap.com

Purpose: To identify real-time trends anywhere in the world on any language supported by Twitter right down to city level



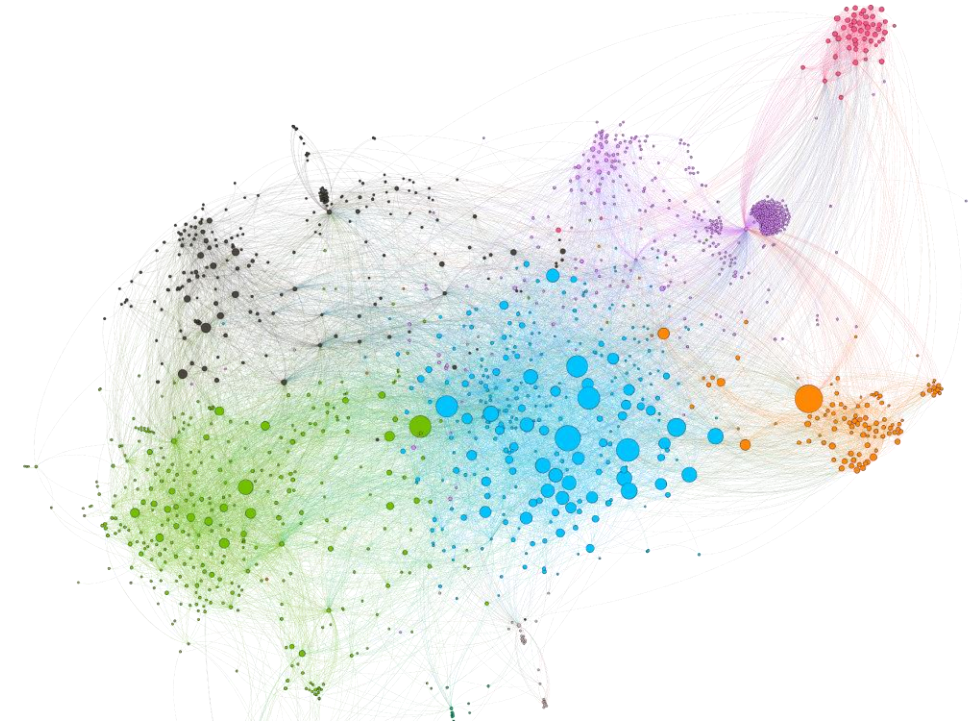


5 - NETWORK VISUALIZATION

SOCIAL NETWORK ANALYSIS (SNA) 101

Social network analysis (SNA) is a strategy for investigating social structures through the use of network and graph theories. It characterizes networked structures in terms of nodes (individual actors, people, pages or things within the network) and edges (relationships or interactions) that connect them.

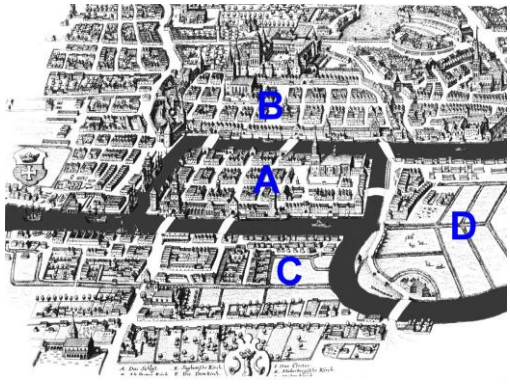
A **sociogram** is the visual representation of these nodes and their relationships (edges) to one another. The colour represents community affiliation and the size of each node represents importance of place in the network.



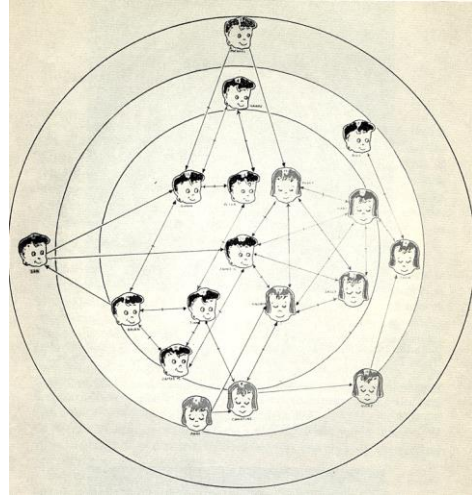
	Reading a Twitter Account Sociogram	Reading a Facebook Page-Like Network Sociogram
Node (circle)	Each node is a Twitter account.	Each node is a Facebook page.
Edge (line)	Each edge is a connection between accounts that interacted at least once with one another via a tweet. Thicker edges mean more tweets/interactions between those accounts.	Each edge is a page like. Pages in Facebook can like other pages. There is no change in edge thickness on these sociograms since pages can only like another page once.
Colour	Accounts that have something in common are generally clustered together and have the same colour due to their interactions.	Accounts that have something in common are generally clustered together and have the same colour due to their interactions.
Size	Larger accounts tend to have more incoming tweets from a greater number of important accounts. The number of followers that an account has does not influence this number.	Larger accounts tend to be liked by a greater number of important pages. The number of individuals who follow a page does not influence this number.

SNA IS NOT NEW

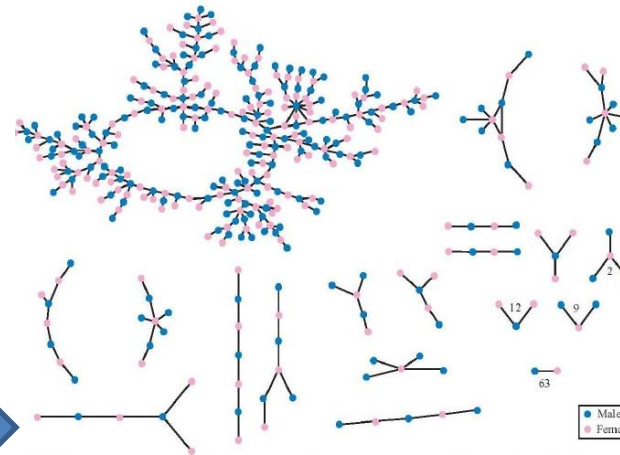
SAME PROCESS BIGGER DATA SETS



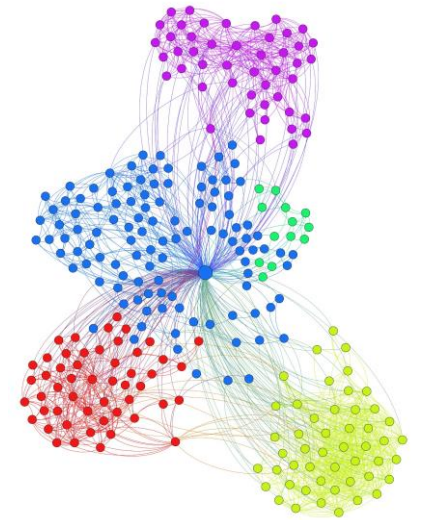
Konigsberg Bridge Problem (1736)



First Grade Class (Grant - 1952)

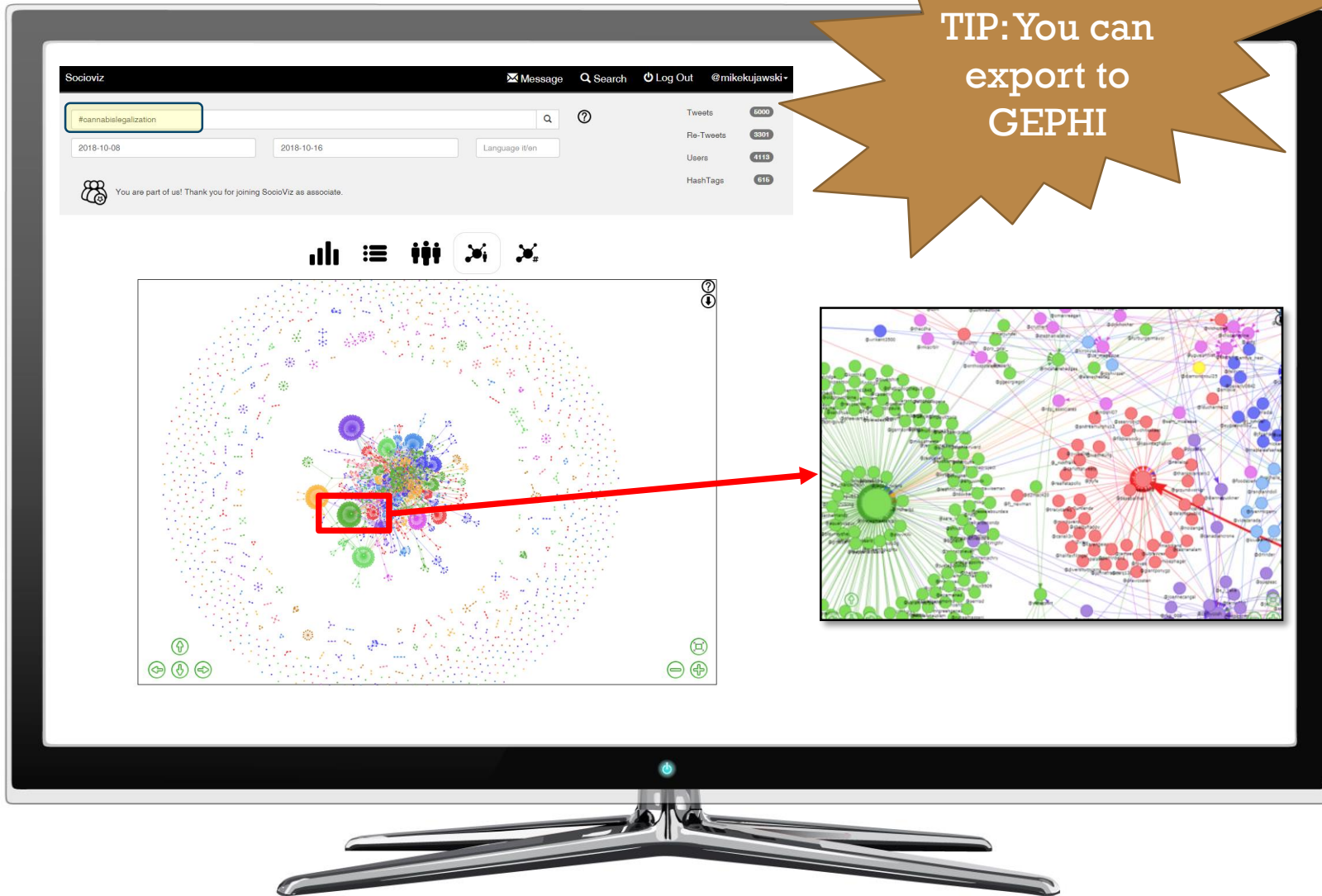


Jefferson High School Sexual Relationship Study (Stovel, Moody, Bearman, 2005)



Facebook Sociogram (2012)

Network analysis has historically been used to study friendship and acquaintance networks, kinship, disease transmission, and sexual relationships



TIP: You can export to GEPHI

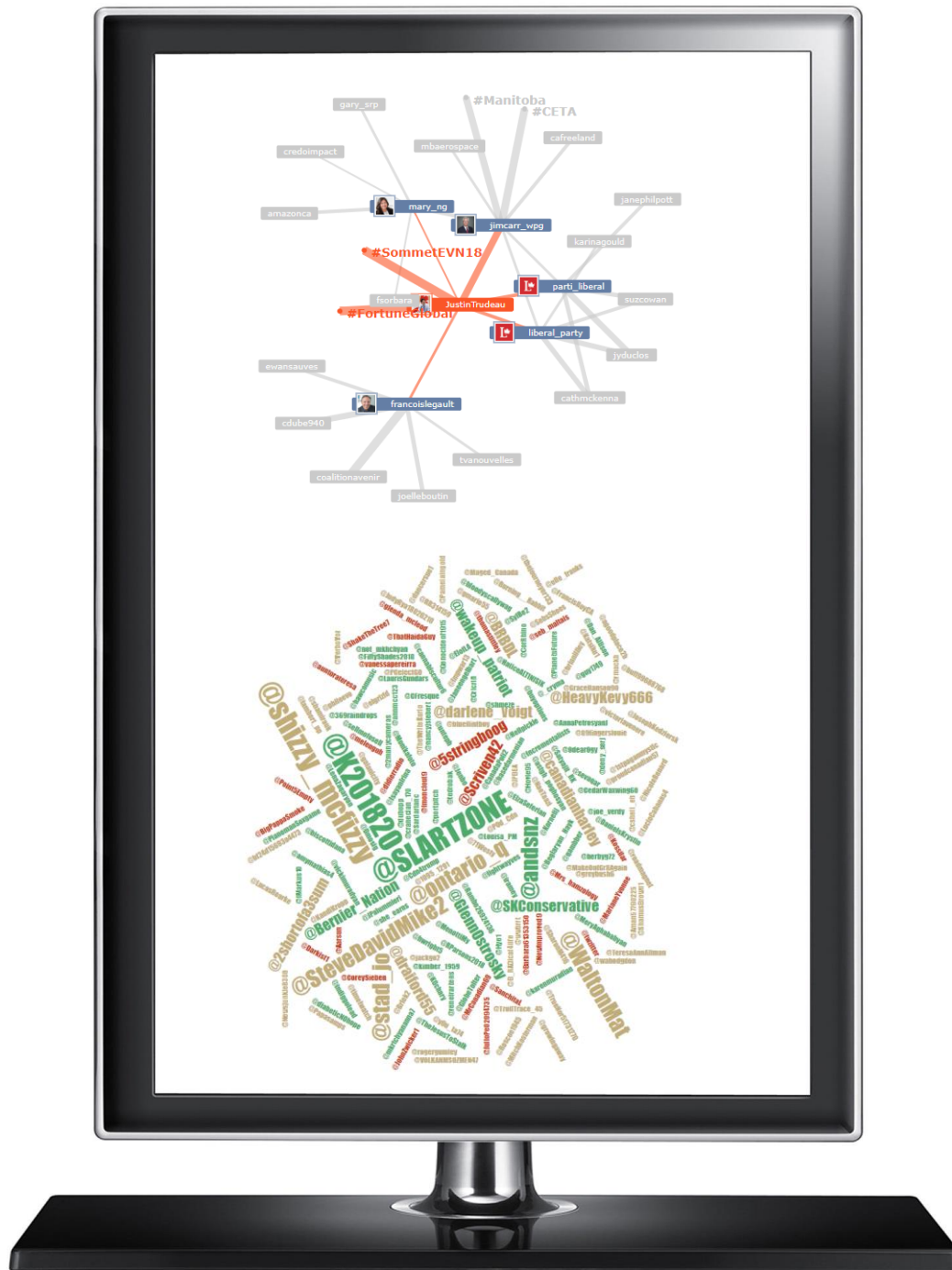
USER & HASHTAG NETWORK VISUALIZATION

socioviz.net

*Free version analyzes 500 of the latest tweets. If you make a donation it allows 5000

Purpose: To quickly identify the key influencers and communities within a topic area + to identify key hashtag sub-topics



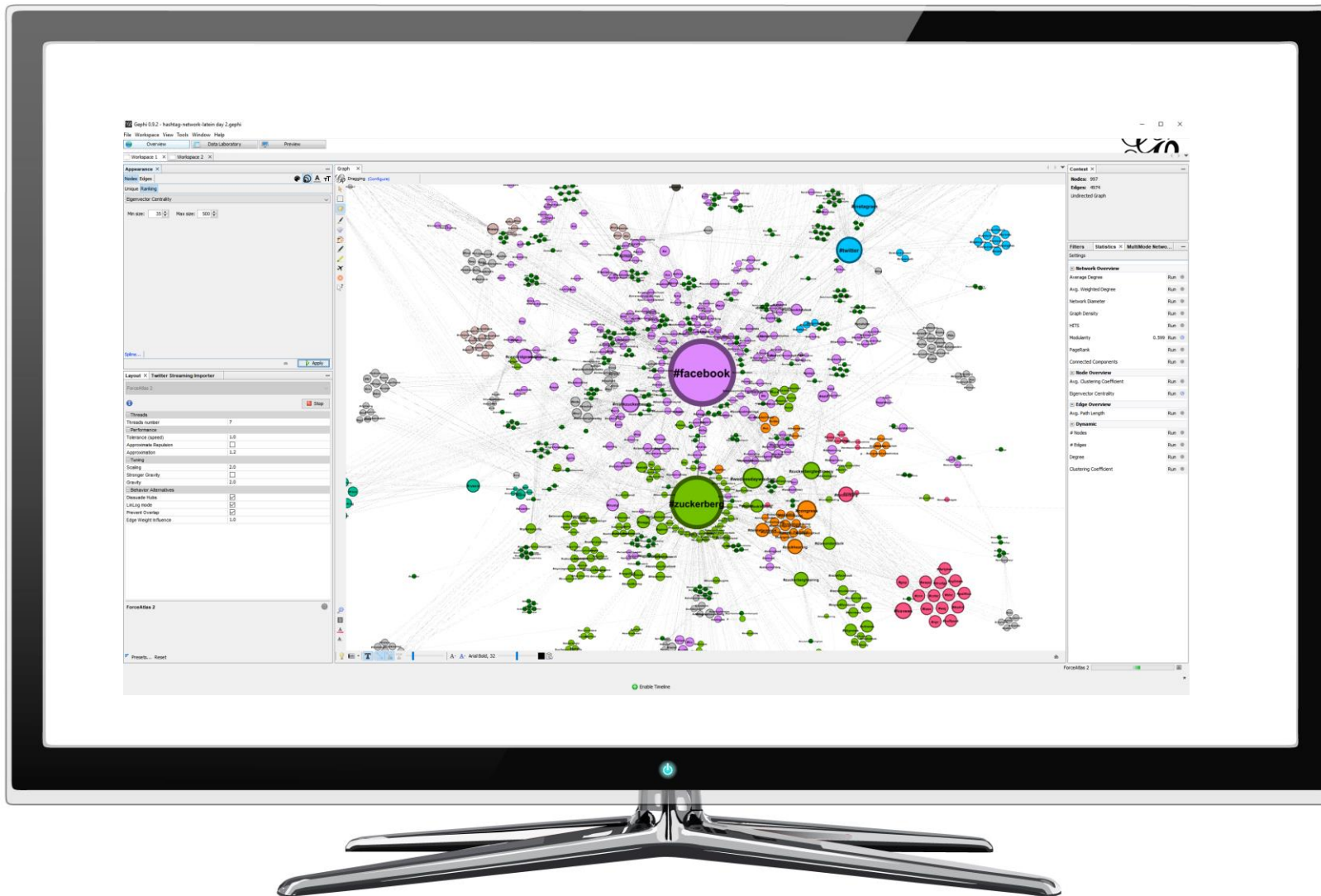


TWITTER NETWORK EXPLORATION

analytics.mentionmapp.com

Purpose: To quickly identify the general topics being discussed around a central user or hashtag + to explore their surrounding networks





ADVANCED SOCIAL NETWORK ANALYSIS

gephi.org

Purpose: This is the leading visualization and exploration software for all kinds of graphs and networks. Best of all it's open source (FREE)



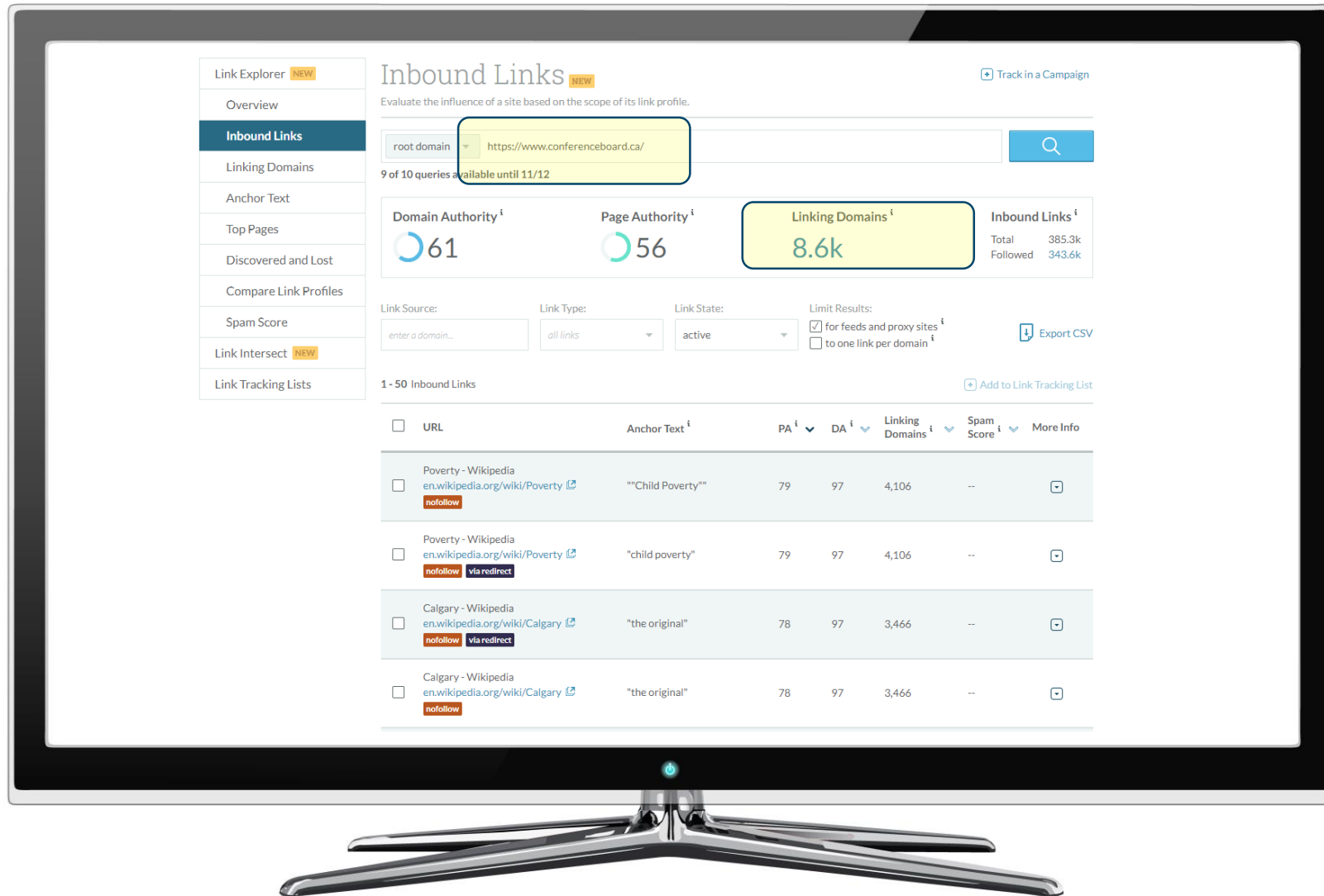


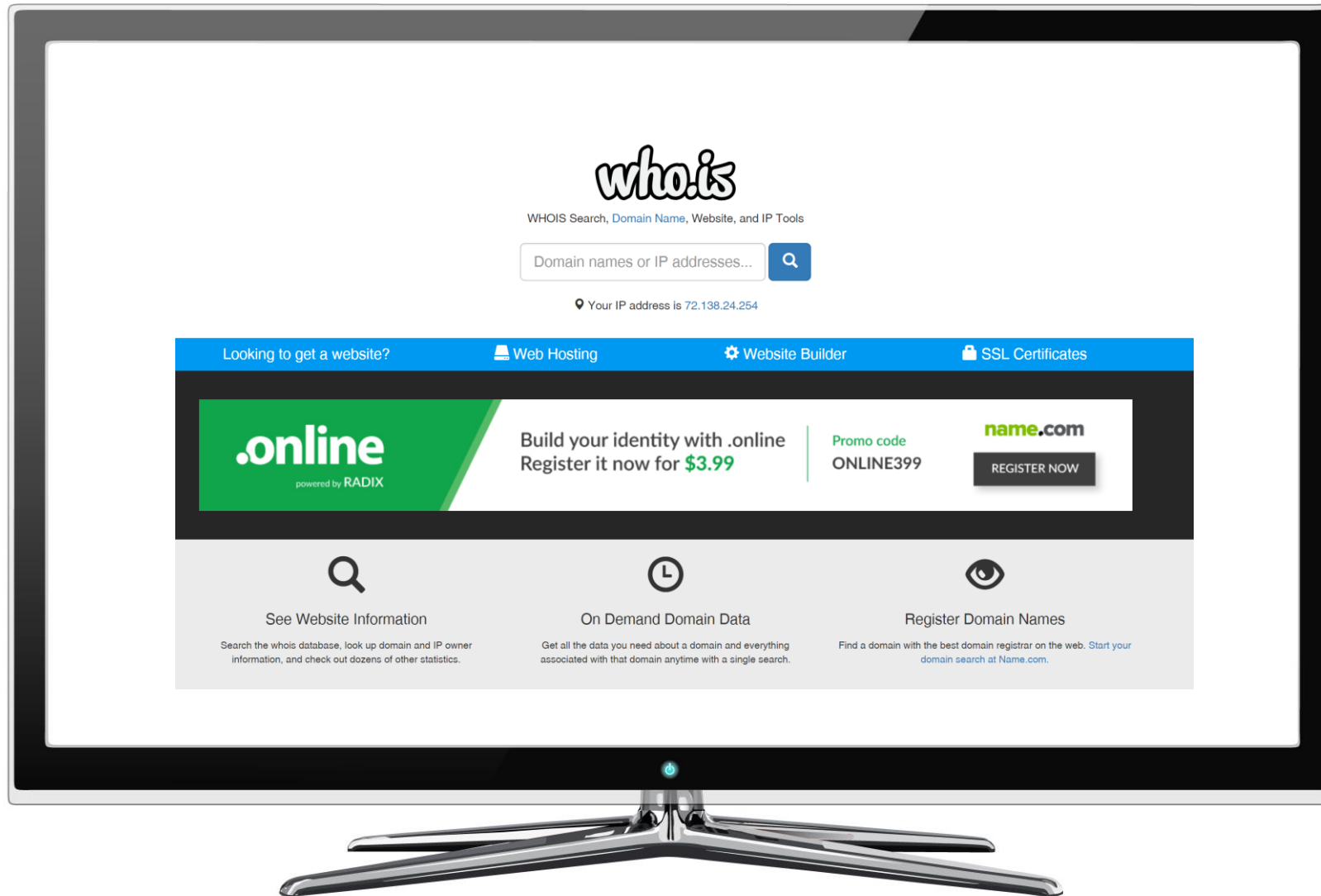
6 - OTHER

INBOUND LINK CHECKING

analytics.moz.com/pro/link-explorer/

Purpose: Determine what other web domains are linking to the web property in question





DOMAIN OWNER RESEARCH

who.is

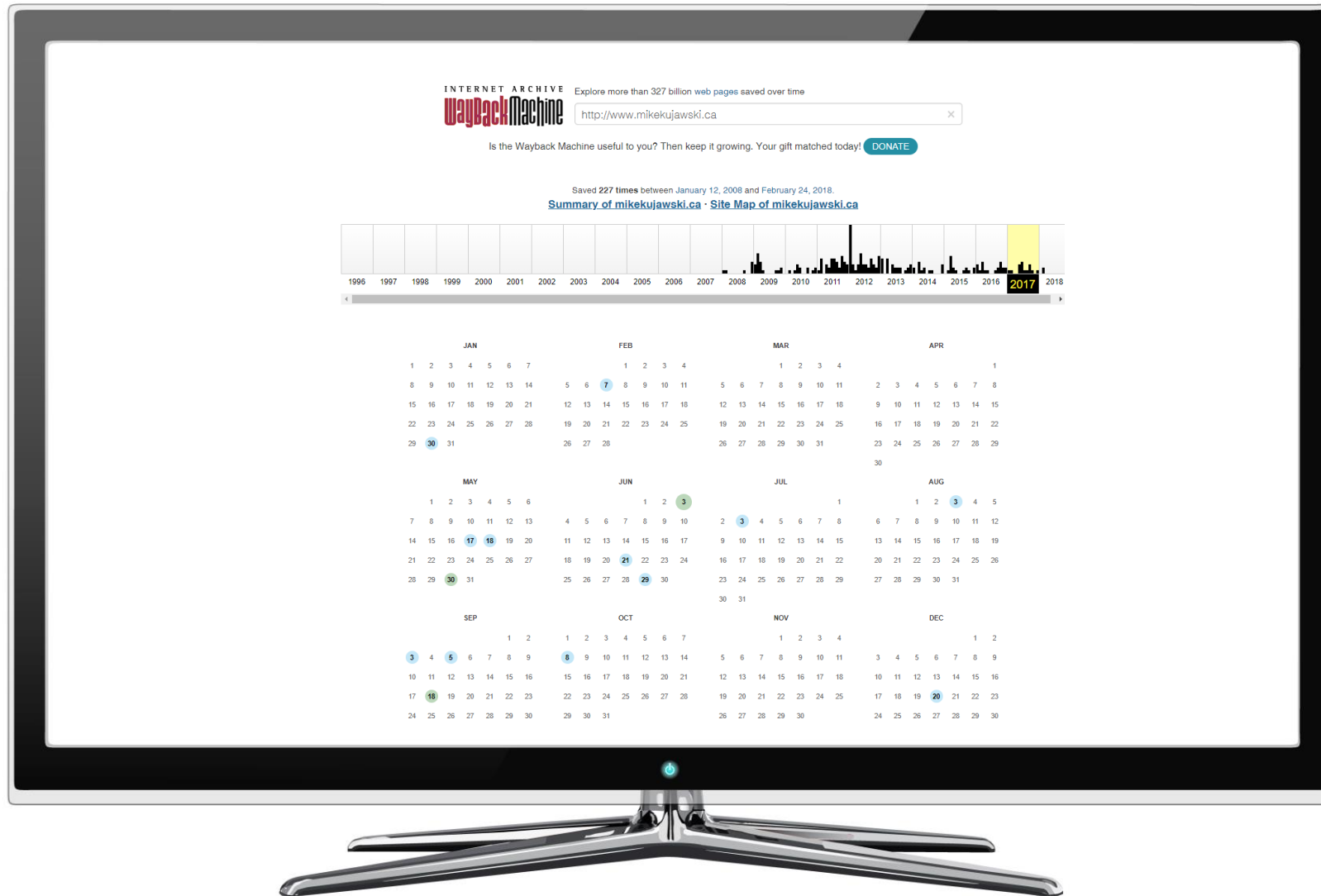
Purpose: Determine who has registered a particular domain. Useful for finding hidden contact info.

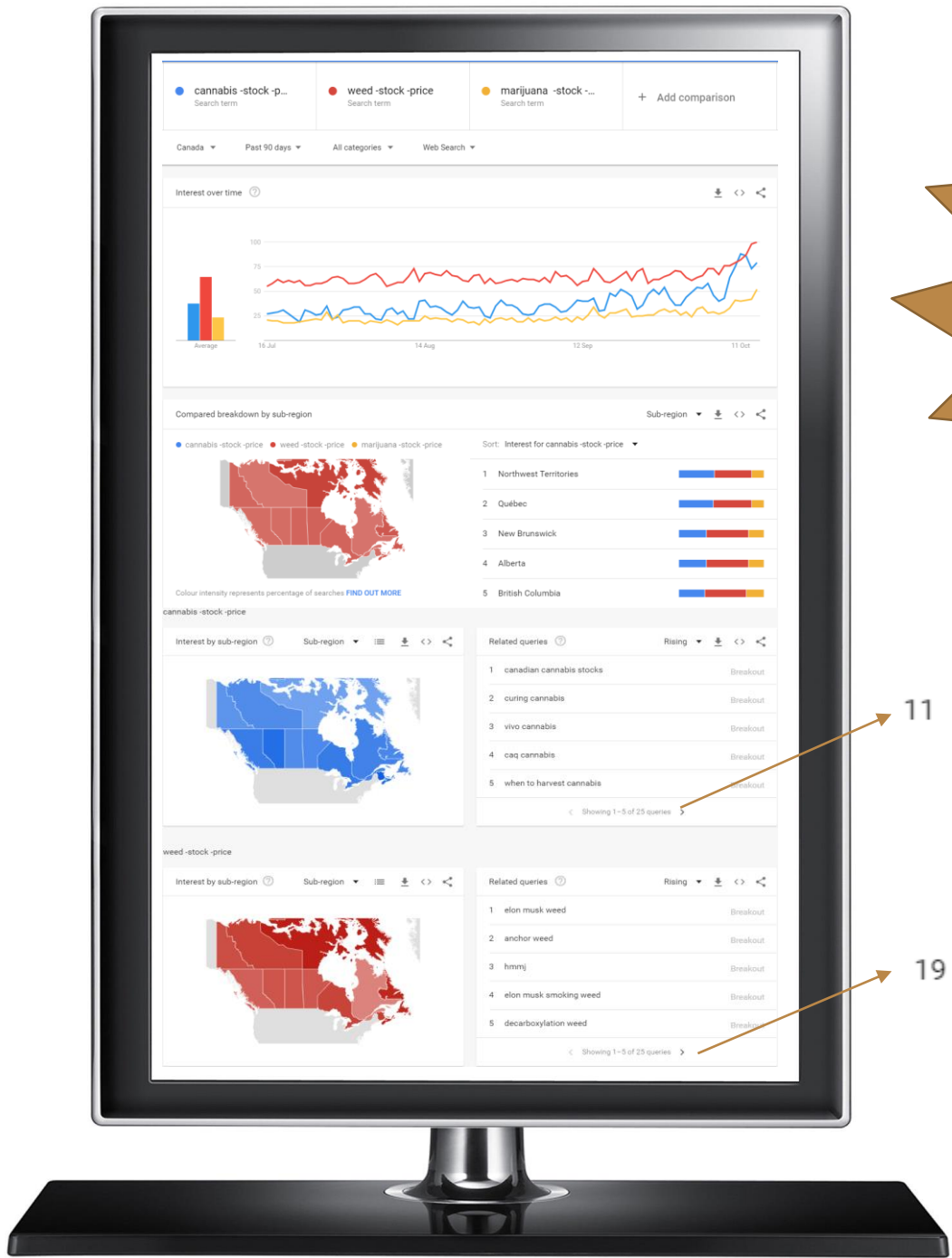


TRAVELING BACK IN TIME

<http://archive.org/web/>

Purpose: Finding old and/or deleted web properties





TIP: You can export to Excel + add exclusion terms

SEARCH TREND DATA

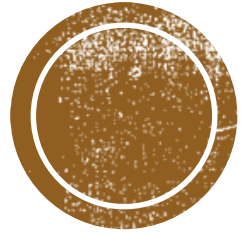
trends.google.com/trends/explore

Purpose: To determine how and when people search for a specific topic

11 how long does weed stay in your pee

19 can you smoke weed when pregnant





THAT'S IT FOR TODAY!

The slides for this presentation will be made available for download.





Thanks

Thank you

Merci

Danke

Dziękuję

Grazie

Mahalo

Takk

Dékoju

Gracias

Asante

Spasibo

Дзякуй

Qagaasakuq

Paldies

Hvala

Hohou

Köszí

Tānan

Many thanks

Asante

Хвала

Faleminderit

Dékuju

Merci

Qujanaq

Danke schön

Tack

Muchas gracias

Terima kasih

Ngiyabonga

Obrigado

Māuruuru

Merci beaucoup

謝謝

Köszí

多謝

Paхмат

Hvala

Vielen Dank

Teşekkürle

EUXαριστώ

Dankon

Mersi

Pilámaya

Ahxéhee

Néá'eše

Кіітос

Тack

Danke schön

Дank u

감사합니다

Спасибо

Муццумецк

Mälö

धन्यवाद





Mike Kujawski

Mobile: 613.899.1348

E-mail: mikekujawski@cepsm.ca

Website: cepsm.ca

LinkedIn: "Mike Kujawski"

Twitter: [@mikekujawski](https://twitter.com/mikekujawski)

Skype: [mikekujawski](https://www.skype.com/people/mikekujawski)

Blog: mikekujawski.ca