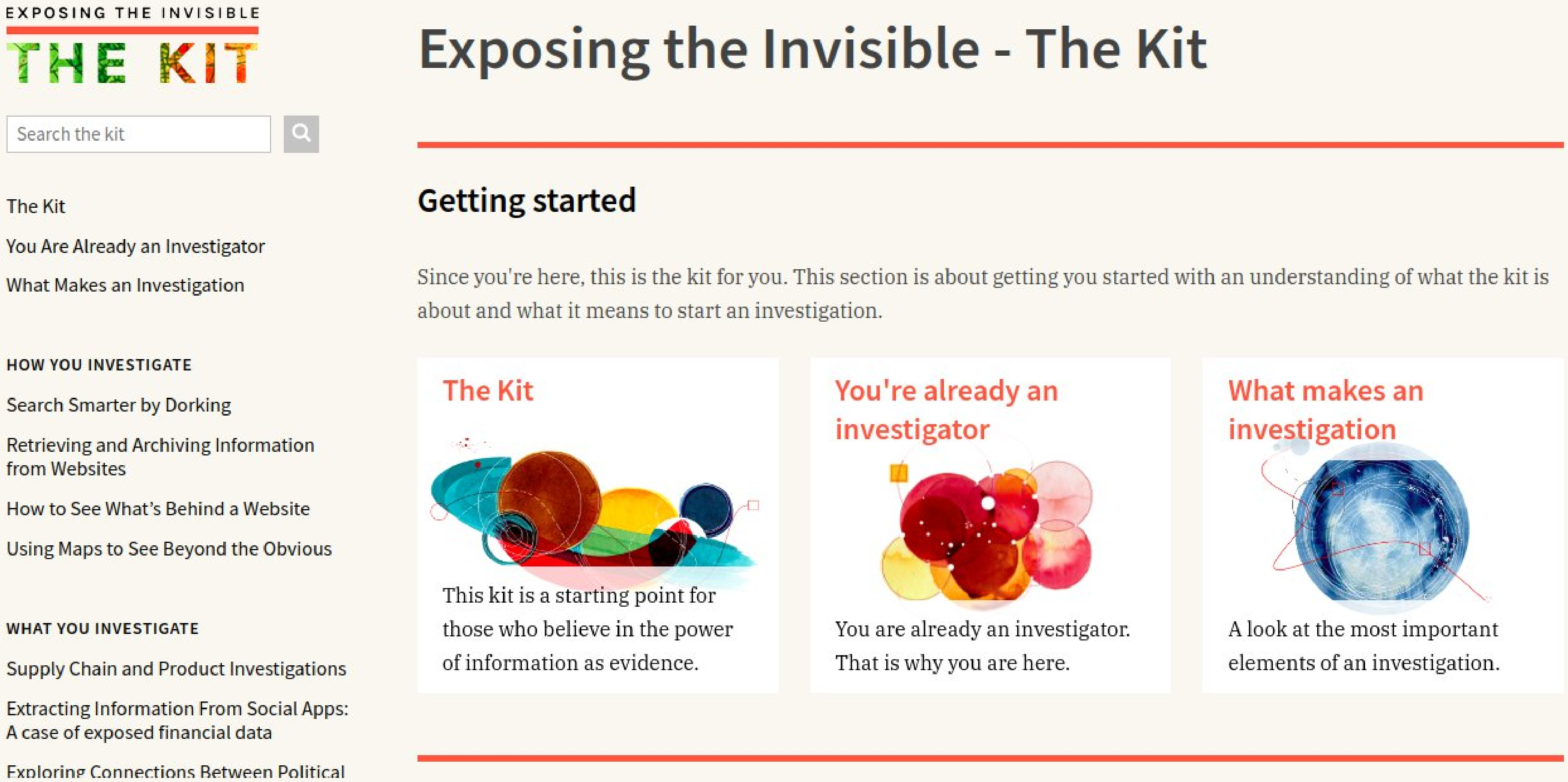
TACTICAL TECH

WHO is behind websites?

20 May 2021 – Data Harvest

Laura Ranca, Tactical Tech - “Exposing the Invisible”
laura@tacticaltech.org
https://kit.exposingtheinvisible.org/
https://exposingtheinvisible.org/
https://tacticaltech.org/

Guides for Investigators: https://kit.exposingtheinvisible.org/



EXPOSING THE INVISIBLE

THE KIT

Search the kit

The Kit

You Are Already an Investigator

What Makes an Investigation

HOW YOU INVESTIGATE

Search Smarter by Dorking

Retrieving and Archiving Information from Websites

How to See What's Behind a Website

Using Maps to See Beyond the Obvious

WHAT YOU INVESTIGATE

Supply Chain and Product Investigations

Extracting Information From Social Apps: A case of exposed financial data

Exploring Connections Between Political

Exposing the Invisible - The Kit

Getting started

Since you're here, this is the kit for you. This section is about getting you started with an understanding of what the kit is about and what it means to start an investigation.

The Kit

This kit is a starting point for those who believe in the power of information as evidence.

You're already an investigator

You are already an investigator. That is why you are here.

What makes an investigation

A look at the most important elements of an investigation.

Outline

- What is WHOIS & why we care
- Research methods and tools
- Safety first
- Resources

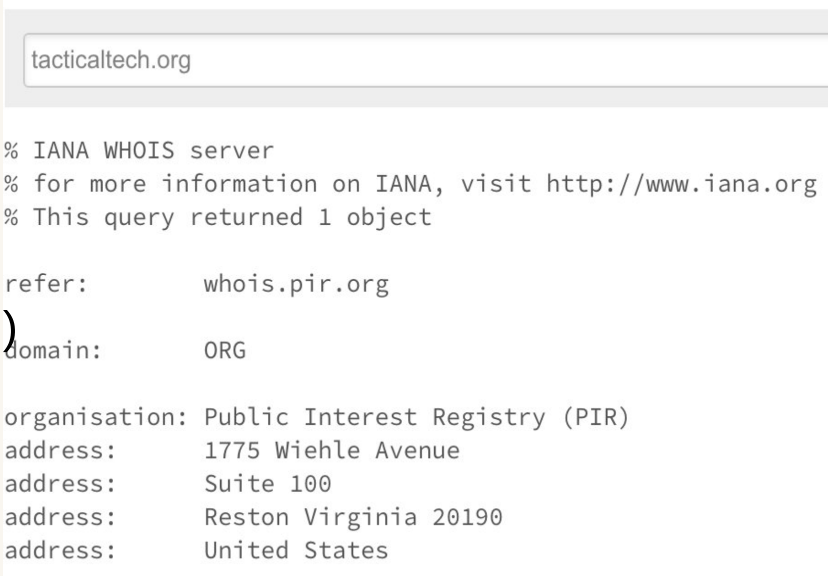
What is

- WHOIS
- Domain name
- Registrar & Registrant
- IP address
- Server

WHOIS searches

finding website ownership / connections

- https://who.is
- https://iana.org/whois
- https://www.whois.com/whois/
- https://godaddy.com/whois
- https://whois.domaintools.com (limited free search)
- https://lookup.icann.org/lookup



tacticaltech.org

```
% IANA WHOIS server
% for more information on IANA, visit http://www.iana.org
% This query returned 1 object

refer:        whois.pir.org

domain:       ORG

organisation: Public Interest Registry (PIR)
address:      1775 Wiehle Avenue
address:      Suite 100
address:      Reston Virginia 20190
address:      United States
```

Tip: Search them all and more, they are endless!

!GDPR has caused a lot of changes in terms of access to WHOIS data! ...but don’t give up.

(see complete guide: https://kit.exposingtheinvisible.org/en/how/web.html)

WHOIS searches

website ownership / connections

Reverse WHOIS search to identify ownership (search by name/email):

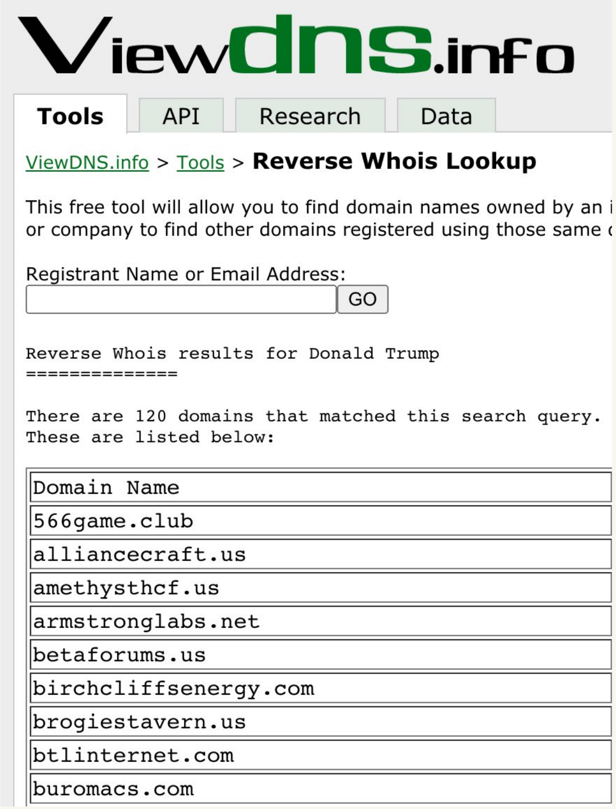
- https://viewdns.info/reversewhois/
- https://domaineye.com/reverse-whois/
- https://reversewhois.domaintools.com/

IP checks:

- https://ipinfo.info/html/ip_checker.php

Reverse IP Search:

- https://viewdns.info/reverseip/



ViewDNS.info > Tools > **Reverse Whois Lookup**

This free tool will allow you to find domain names owned by an i
or company to find other domains registered using those same

Registrant Name or Email Address:

```
Reverse Whois results for Donald Trump
==============

There are 120 domains that matched this search query.
These are listed below:
```

Domain Name
566game.club
alliancecraft.us
amethysthcf.us
armstronglabs.net
betaforums.us
birchcliffsenergy.com
brogiestavern.us
btlinternet.com
buromacs.com

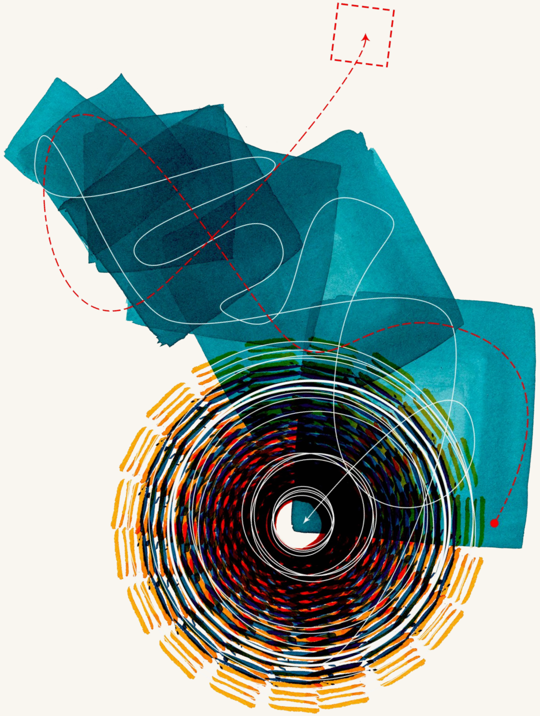
Website data indexing

- **Robots.txt** (tells a crawler what not to index)

  *eg. *https://site.org/robots.txt*

- **Sitemap.xml** (tells a crawler what to index)

  *eg: *https://site.org/sitemap.xml*



*(Where *https://site.org is any website you wish to check.)*

Using Web Archives to Find Connections

Web Archives

preserving and recovering online content

- Website archives: collecting history
- Archiving websites: making history *(social media and images/videos are difficult to archive and recover so take additional measures to preserve them)*

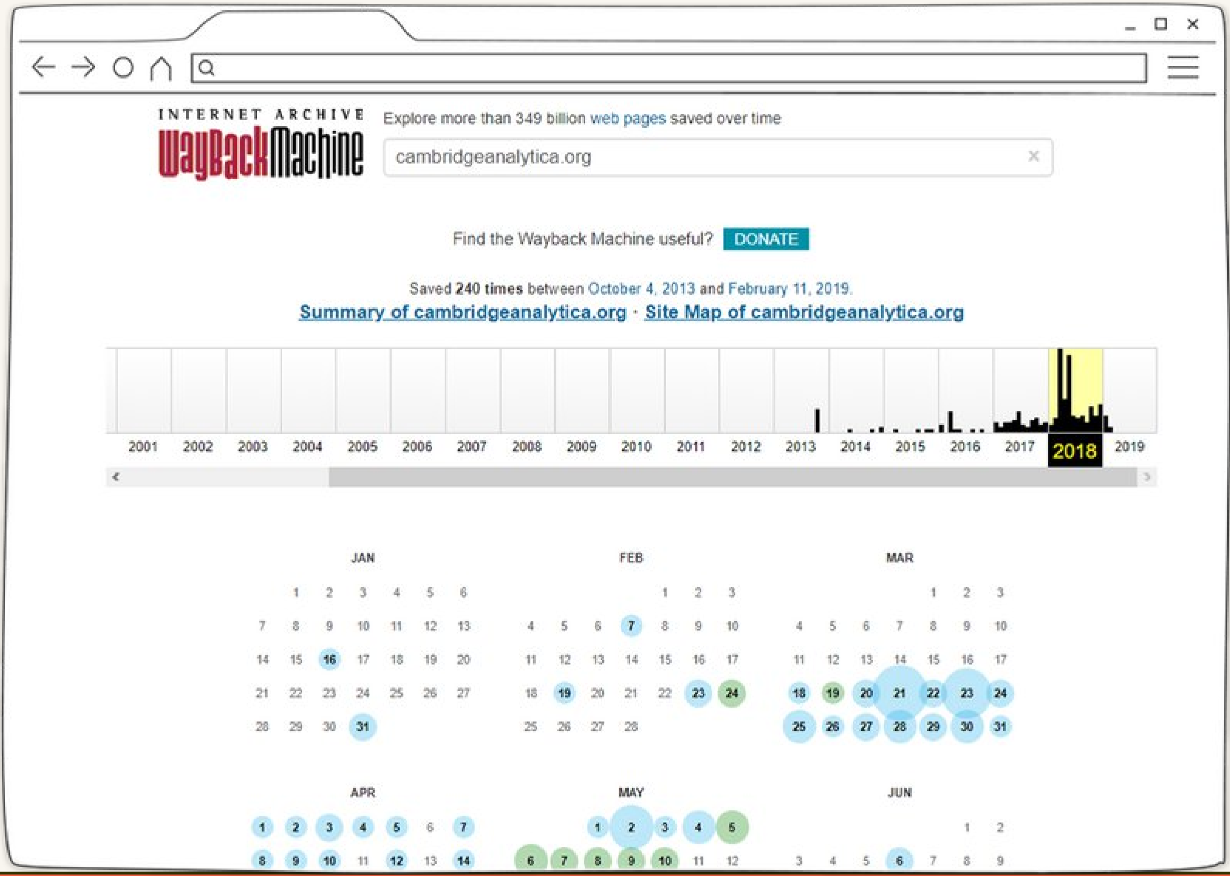
WaybackMachine: https://archive.org/web/ [redacted]

Archive Today: https://archive.vn/ [redacted]

(see complete guide: https://kit.exposingtheinvisible.org/en/how/web-archive.html)

The Wayback Machine: https://archive.org/web/

Search for a website's archive: e.g. *www.cambridgeanalytica.org*



INTERNET ARCHIVE WaybackMachine

Explore more than 349 billion web pages saved over time

cambridgeanalytica.org

Find the Wayback Machine useful? DONATE

Saved 240 times between October 4, 2013 and February 11, 2019.

Summary of cambridgeanalytica.org · Site Map of cambridgeanalytica.org

[redacted]

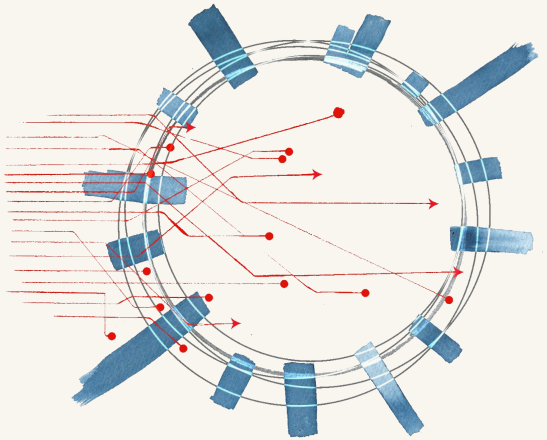
The Wayback Machine:

browser searches for archived sites

- https://web.archive.org/*www.yoursite.com/* - to reach the homepage of the website on Wayback Machine

- https://web.archive.org*/*/www.yoursite.com/* - to obtain a calendar view of the website on Wayback Machine

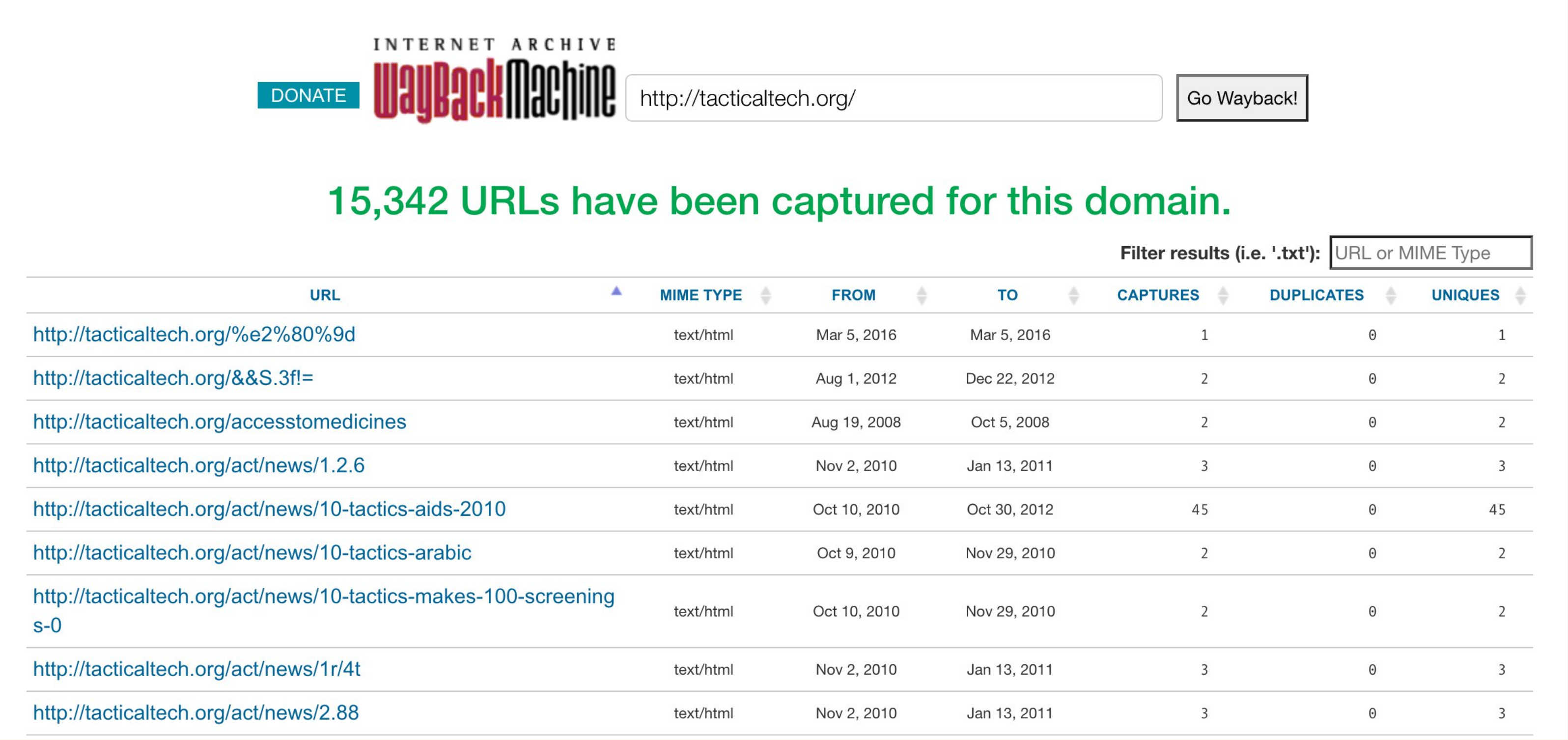
- https://web.archive.org*/*/www.yoursite.com/** - to search for all pages archived not just homepage



[redacted]

Example search for https://web.archive.org/web/*/tacticaltech.org/*

(https://web.archive.org/*/*www.yoursite.com/**)



15,342 URLs have been captured for this domain.

Filter results (i.e. '.txt'): URL or MIME Type

URL	MIME TYPE	FROM	TO	CAPTURES	DUPLICATES	UNIQUES
http://tacticaltech.org/%e2%80%9d	text/html	Mar 5, 2016	Mar 5, 2016	1	0	1
http://tacticaltech.org/&&S.3f!=	text/html	Aug 1, 2012	Dec 22, 2012	2	0	2
http://tacticaltech.org/accesstomedicines	text/html	Aug 19, 2008	Oct 5, 2008	2	0	2
http://tacticaltech.org/act/news/1.2.6	text/html	Nov 2, 2010	Jan 13, 2011	3	0	3
http://tacticaltech.org/act/news/10-tactics-aids-2010	text/html	Oct 10, 2010	Oct 30, 2012	45	0	45
http://tacticaltech.org/act/news/10-tactics-arabic	text/html	Oct 9, 2010	Nov 29, 2010	2	0	2
http://tacticaltech.org/act/news/10-tactics-makes-100-screenings-0	text/html	Oct 10, 2010	Nov 29, 2010	2	0	2
http://tacticaltech.org/act/news/1r/4t	text/html	Nov 2, 2010	Jan 13, 2011	3	0	3
http://tacticaltech.org/act/news/2.88	text/html	Nov 2, 2010	Jan 13, 2011	3	0	3

The Wayback Machine: Limitations

- Password-protected websites are not archived.

- Dynamic websites that rely heavily on JavaScript may not be archived properly.

- Website administrators can explicitly request that their sites not be archived, either by publishing a restrictive robots.txt file or by sending a direct request to the Internet Archive.

- Website administrators can request that previously archived content be removed from the Wayback Machine.

[redacted]

- There is currently no full-text search available.

Archive.today: https://archive.vn/

- Allows search of full text of its archives.
- Ignores any restrictions that might be specified in the robots.txt files of the websites that it archives.
- Snapshots of some pages, such as public Facebook profiles and Twitter posts.
- Saves both a text copy and a graphical screenshot of the archived pages

Limitations:

[redacted]

- Does not crawl full websites automatically

Copy Website

HTTrack WEBSITE COPIER

https://www.httrack.com/

Safety First!

Safety First!

‘Fishing’ for data online: risk assessment and mitigation

- Next are optional resources and tips to consult individually
- + Read basic Safety guide for investigators:

https://kit.exposingtheinvisible.org/en/safety.html

Basics – when using online Databases

- Be careful when creating accounts / log-in details
- Use dummy email accounts (not linked to your personal or work details). Generally, separate personal emails from work.
- Use strong passwords – not the same password for different sites, emails, devices, etc.
- Use privacy-respectful browsers

Creating & Maintaining Secure Passwords

- Make it **long** and **random** (hellohowareyou12345 vs.

  Drugge!2$#@droseriotingabsolutely2001@!phewycakes)

- **Not personally** identifiable (no birth date, address, names of your puppies/kids/lovers/haters..)

- Keep it **secret**

- Make it **practical** to remember

- **Unique** – to avoid major damage if it gets exposed

- Keep it **fresh** – change it every now and then

Password Managers

- Allow for strong one-time passwords + just one pass to remember
- Don’t forget the master password
- KeepassXC (offline), LastPass (online), Firefox Lockwise (online)

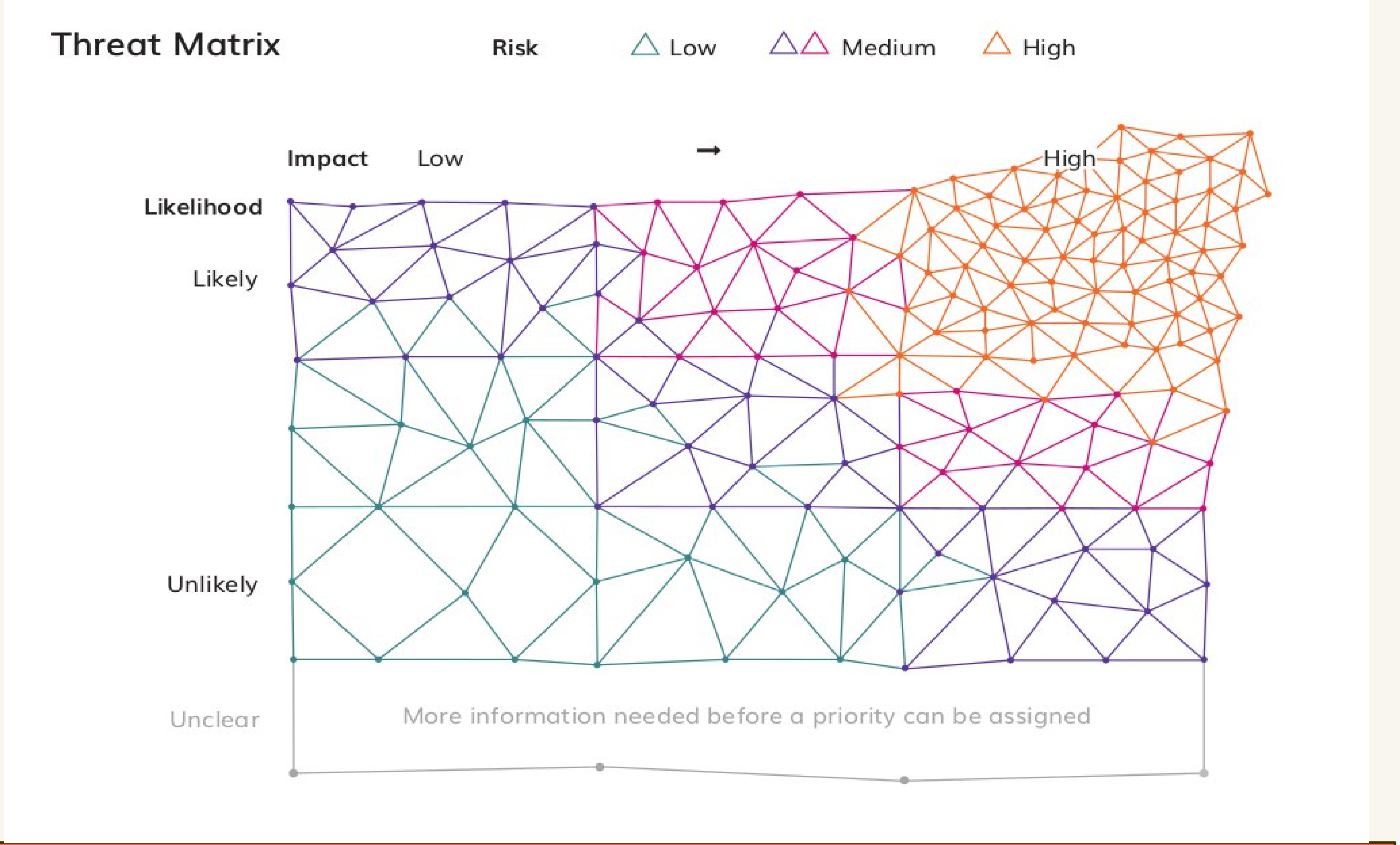
Safety First! - Plan & Actions

- Goal of your research/investigation
- Functions you perform
- Methods and tools you use to perform the functions
- Data you collect/ safeguard/ communicate
- Context(s) where you perform those functions

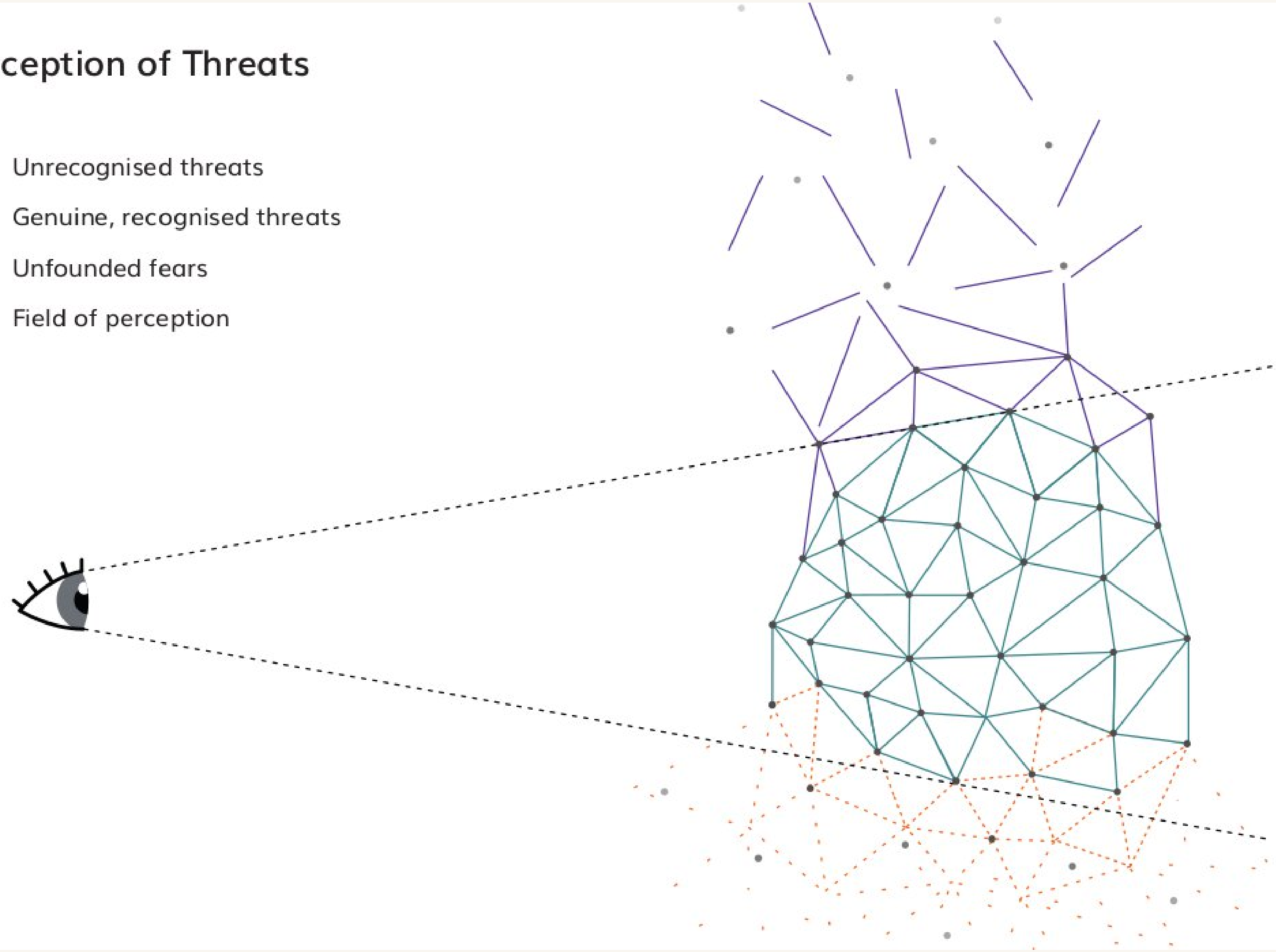
Safety First! - Plan & Actions vs. Threats

- Goal of your research/investigation
- Functions you perform
- Methods and tools you use to perform the functions
- Data you collect/ safeguard/ communicate
- Context(s) where you perform those functions
  - Adversaries
  - Capabilities of adversaries
  - Consequences

Safety First! - Risk = Threat x Likelihood x Impact



Perception of Threats



△ Unrecognised threats

△ Genuine, recognised threats

△ Unfounded fears

---- Field of perception

Safety First! - Scenarios for Threat Assessment

Scenarios **Threat types**	Tracing businessperson and company connections / e.g. on tax evasion	Tracing politician and private company connections / e.g. on corruption, public money fraud	Tracing organised crime figure and company connections/ e.g. money laundering, trafficking	Tracing company links to abuse /e.g. environmental damage or human rights abuse
A. Physical threat (you and others)				
B. Surveillance (you and others)				
C.Information theft (you and others)				
D. Legal threat (you and others)				
E. Reputation threat (you and others)				

Risk Assessment & Risk Mitigation

- **Risk assessment** - only useful if you are ready to take action based on it.
- **Risk management** - actively preventing risks from happening and mitigating their impact if they happen.
- **Risk management** - purpose is to minimise:
    - minimise likelihood (the possibility of it happening)
    - minimise impact (lower the severity)
    - fix the damage (the effects it causes) of risks.

Example **- Risk Assessment**: online research to build profile of a person and/or company suspected for money laundering connections.

Risk type	Risk	Severity (1 to 5)	Mitigation
Personal	My online accounts and passwords risk being exposed	4	Create dummy accounts for email and database log-ins wherever possible, use KeepsXC or LastPass to manage online account passwords...
Personal and Research goal	Online surveillance, browsing, accounts, data	5	Use TOR and/or VPN for searches, use private browsing windows and privacy aware browsers (eg DuckDuckGo) to erase search history, avoid online transfers and storage of data if possible
Research goal	Online data leaks, access to my cloud files - data is blocked / deleted	5	Keep multiple data back-ups offline at more locations, avoid storing data online if possible.
Sources, others	Risk of identifying who my human sources are	3	If you have a human sources/insider/witness already, don’t start by searching detailed information or clues that only one or a small number of people might hold - you attract attention on them from the start. Start broad, keep the source information secure and secure your searches as above.

Humans are the weakest link

- What methods, services/tools you choose
- What you click (phishing)
- What you choose to share
- How you communicate
- How you act when in the field

Digital - Physical

- Each research/investigation can have a digital and a physical safety aspect and thus related risks to assess.

- Each digital safety aspect can have a risk to your/others' physical safety as well.

Countering Perceptions

- Shift+ Del does not delete your data
- SSD drives do not guarantee deletion
- Forensics retrieve deleted files
- Your finger is not secure enough
- Screen passwords are not a way to secure data
- Incognito mode does not make you anonymous
- Secure tools don’t make you safe

Security considerations in a tool

- Open source
- End to end encrypted
- Does not store data unnecessarily
- Does not leak data
- Does not share data

How we choose tools

1. **Open source**
2. **Trusted** (audited)
3. **Mature** (stable, with an active user-based community and responsive developer community)
4. **User-friendly**
5. **Multi-language** with localisation support (so you can find your own language or localise it)
6. **Multi-platform** (Mac, Windows, Linux, Android)
7. **Documentation** available

Networks

- Firefox private mode, DuckDuckGo, Brave
- Tor Browser
- Trusted Apps
- HTTPS (Security not Privacy)
- VPN

[redacted]

brave

[redacted]

Communication

- Trusted Apps
- Encrypted messaging apps (Wire/Signal)
- PGP
- VPN/Tor

[redacted]

[redacted]

signal

20 years GnuPG

Online Accounts

- Secure your credentials
- Password Manager
- Not sharing extra information
- Two Factor Authentication
- Protect your data from online platforms
- Incognito mode, clear browser history
- Privacy Badger, HTTPS

Privacy Badger

DATA DETOX KIT

https://datadetoxkit.org/

Devices and Data

- Full Disk Encryption: *Bitlocker* (Windows), *FileVault* (Mac), *dm-crypt* (Linux)
- *VeraCrypt*, *Cryptomator* file containers for data
- Backup your data in case of loss of devices (*Duplicati, Clonezilla, Spider Oak, NextCloud, Tresorit, Google**...and hard drives)

Cryptomator

tresorit

Your Digital Safety

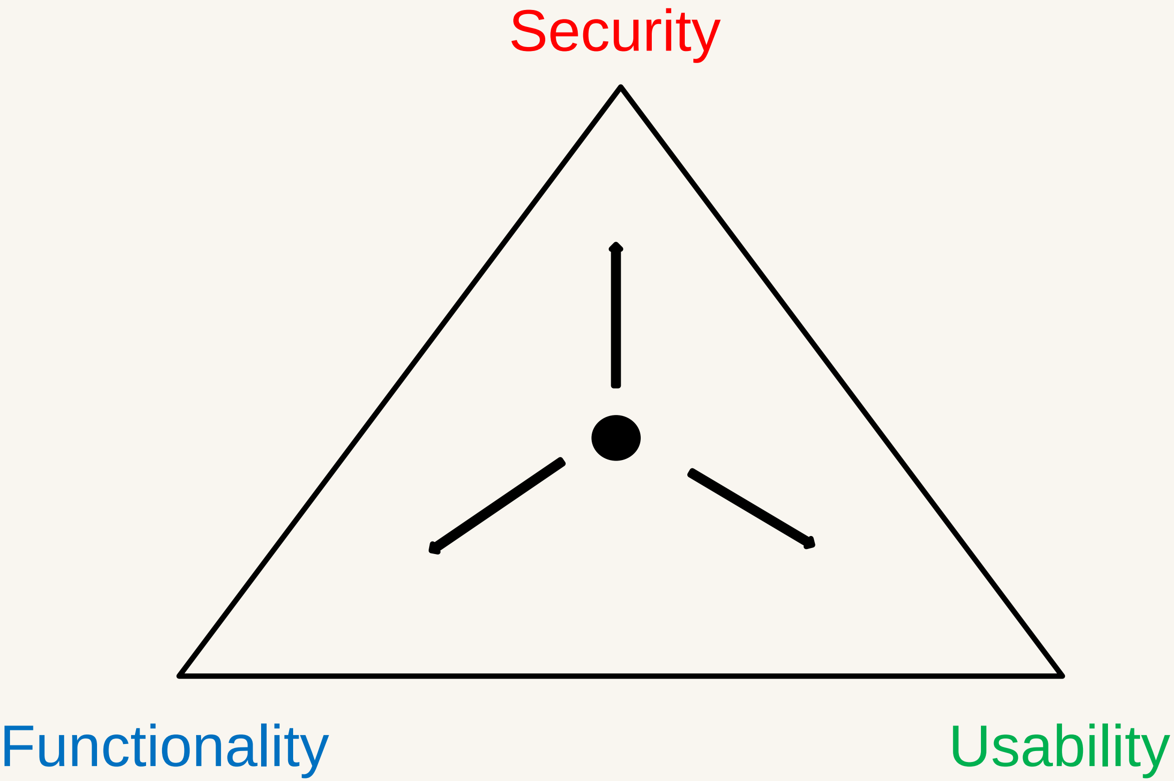
□ is about the function you perform and the context in which the function is performed

□ cannot be approached in isolation from overall safety

□ cannot be approached in isolation from that of other people you communicate with

(see basic Safety First guide: https://kit.exposingtheinvisible.org/en/safety.html)

Digital Safety Trade-off



Should I always go for most secure tool?

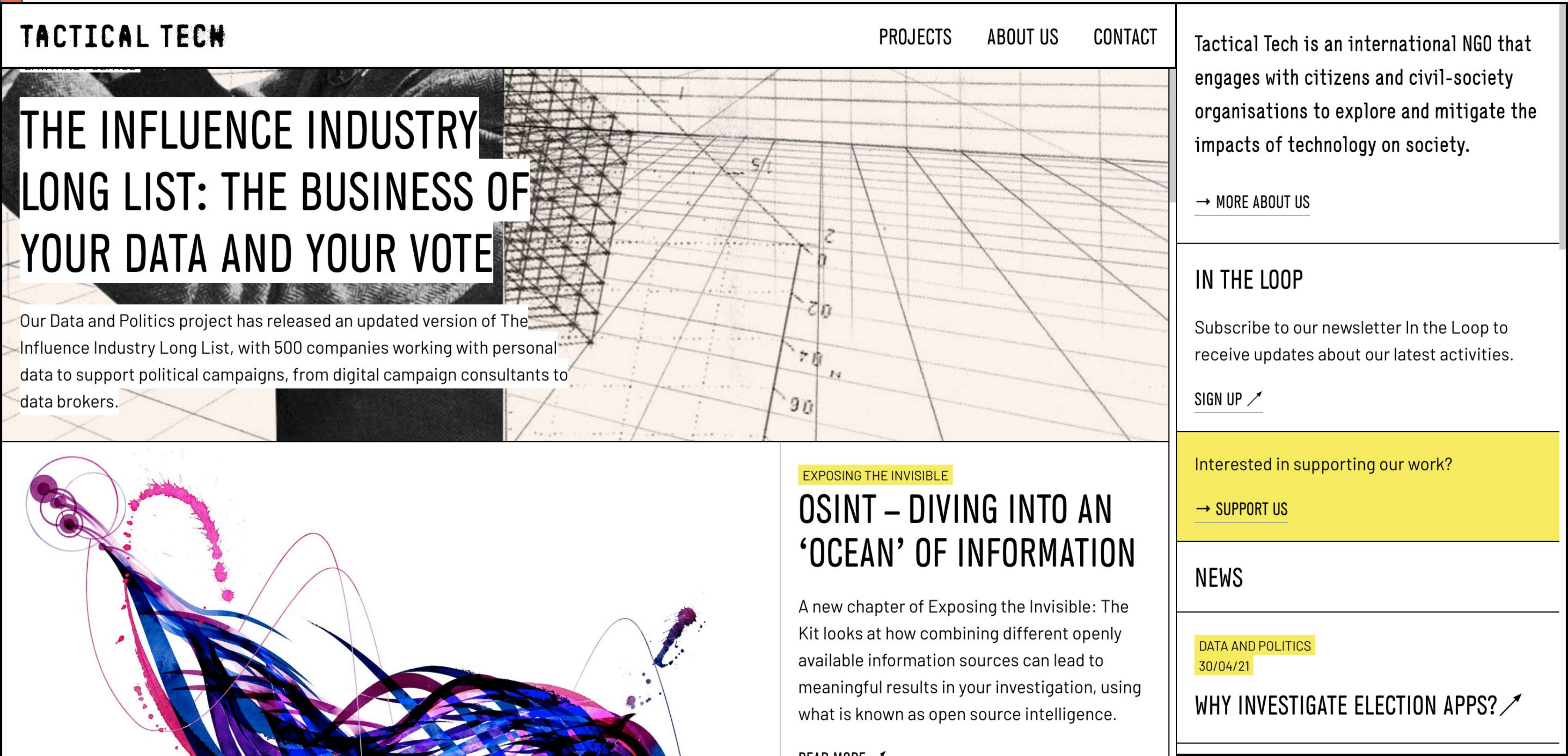
NO

- Secure tools don’t mean security
- Some security tools draw more attention (depending on country)
- Harder to adopt with no added value depending on context

Resources for Research

- Repository of tools and indexed website ownership/connections research: https://osint.sh/ (recently launched)
- Basic WHOIS guide + cases: https://kit.exposingtheinvisible.org/en/how/web.html
- Blog about OSINT tools and workflows: https://jakecreps.com/tag/osint-tools/
- Tips and tutorials: https://osintcurio.us/10-minute-tips/

About Tactical Tech and our projects: https://tacticaltech.org/



TACTICAL TECH

PROJECTS ABOUT US CONTACT

THE INFLUENCE INDUSTRY LONG LIST: THE BUSINESS OF YOUR DATA AND YOUR VOTE

Our Data and Politics project has released an updated version of The Influence Industry Long List, with 500 companies working with personal data to support political campaigns, from digital campaign consultants to data brokers.

EXPOSING THE INVISIBLE

OSINT – DIVING INTO AN 'OCEAN' OF INFORMATION

A new chapter of Exposing the Invisible: The Kit looks at how combining different openly available information sources can lead to meaningful results in your investigation, using what is known as open source intelligence.

READ MORE ↗

Tactical Tech is an international NGO that engages with citizens and civil-society organisations to explore and mitigate the impacts of technology on society.

→ MORE ABOUT US

IN THE LOOP

Subscribe to our newsletter In the Loop to receive updates about our latest activities.

SIGN UP ↗

Interested in supporting our work?

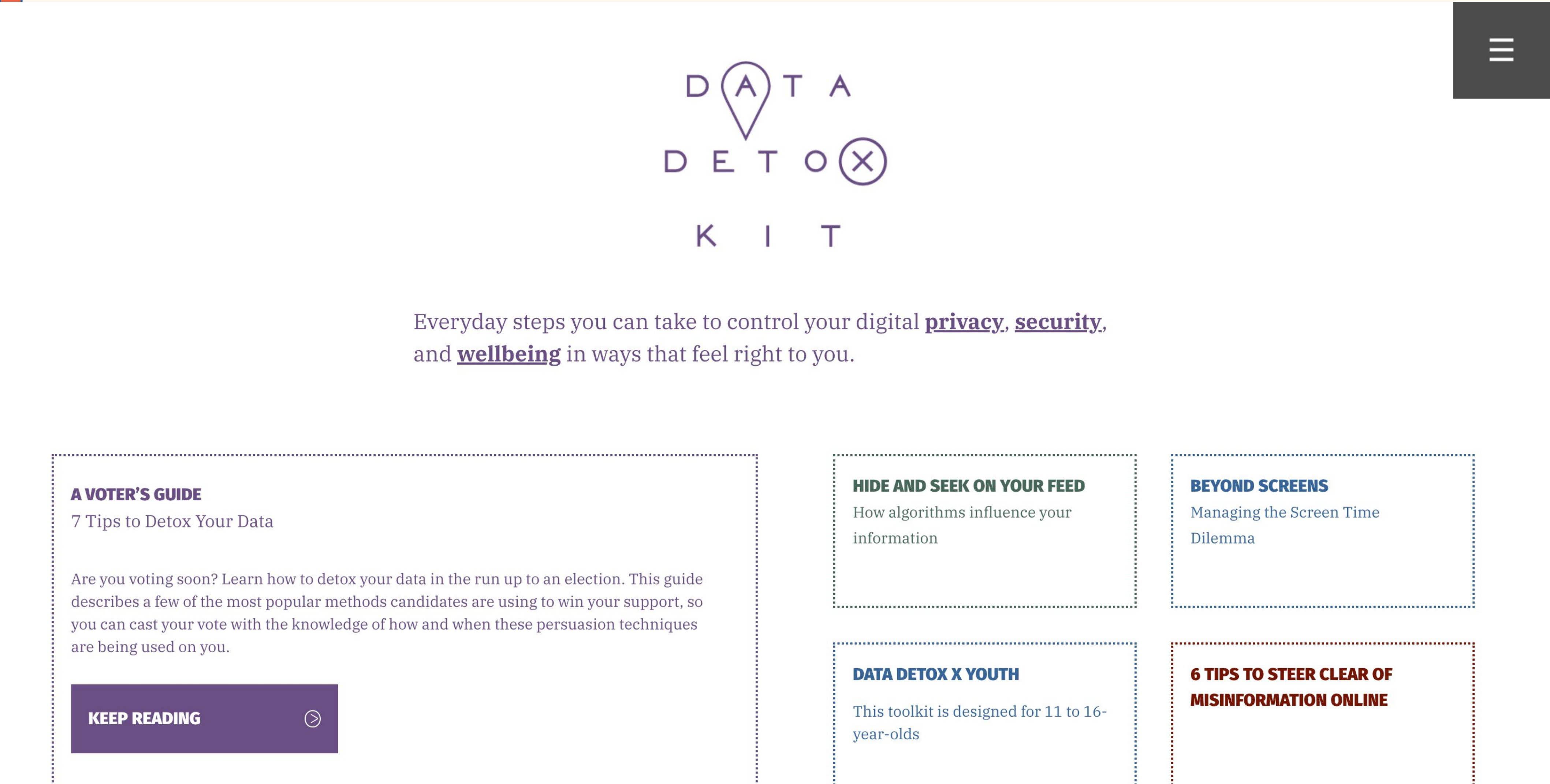
→ SUPPORT US

NEWS

DATA AND POLITICS
30/04/21

WHY INVESTIGATE ELECTION APPS? ↗

Detoxify your Digital Self: https://datadetoxkit.org/



DATA DETOX KIT

Everyday steps you can take to control your digital **privacy**, **security**, and **wellbeing** in ways that feel right to you.

A VOTER'S GUIDE

7 Tips to Detox Your Data

Are you voting soon? Learn how to detox your data in the run up to an election. This guide describes a few of the most popular methods candidates are using to win your support, so you can cast your vote with the knowledge of how and when these persuasion techniques are being used on you.

HIDE AND SEEK ON YOUR FEED

How algorithms influence your information

BEYOND SCREENS

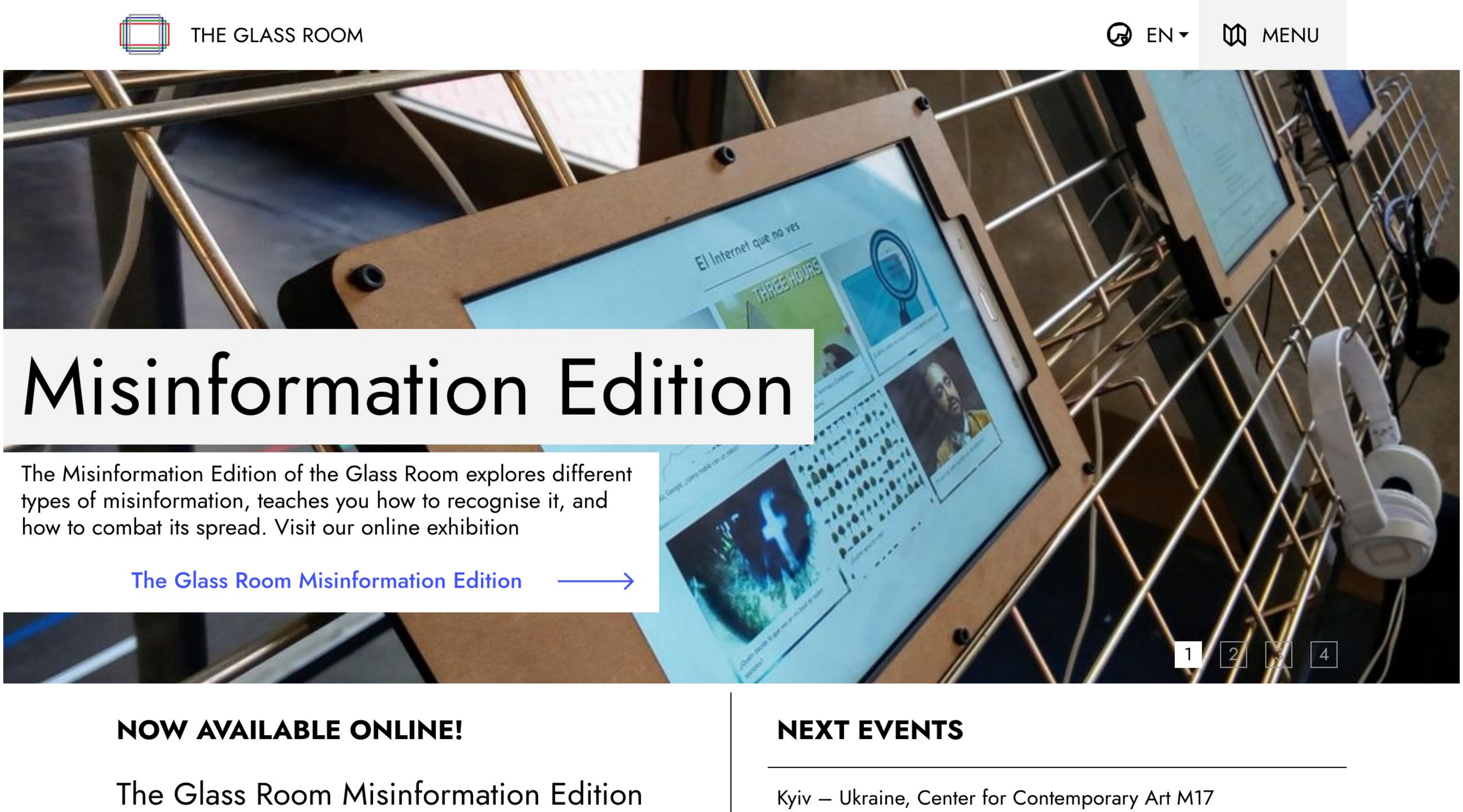
Managing the Screen Time Dilemma

DATA DETOX X YOUTH

This toolkit is designed for 11 to 16-year-olds

6 TIPS TO STEER CLEAR OF MISINFORMATION ONLINE

Online exhibition on tech and privacy https://theglassroom.org/



Misinformation Edition

The Misinformation Edition of the Glass Room explores different types of misinformation, teaches you how to recognise it, and how to combat its spread. Visit our online exhibition

The Glass Room Misinformation Edition →

NOW AVAILABLE ONLINE!

The Glass Room Misinformation Edition

NEXT EVENTS

Kyiv – Ukraine, Center for Contemporary Art M17

Thank you!

Laura @ tacticaltech.org

Exposing the Invisible
https://exposingtheinvisible.org/
https://kit.exposingtheinvisible.org/

Tactical Tech
https://tacticaltech.org/