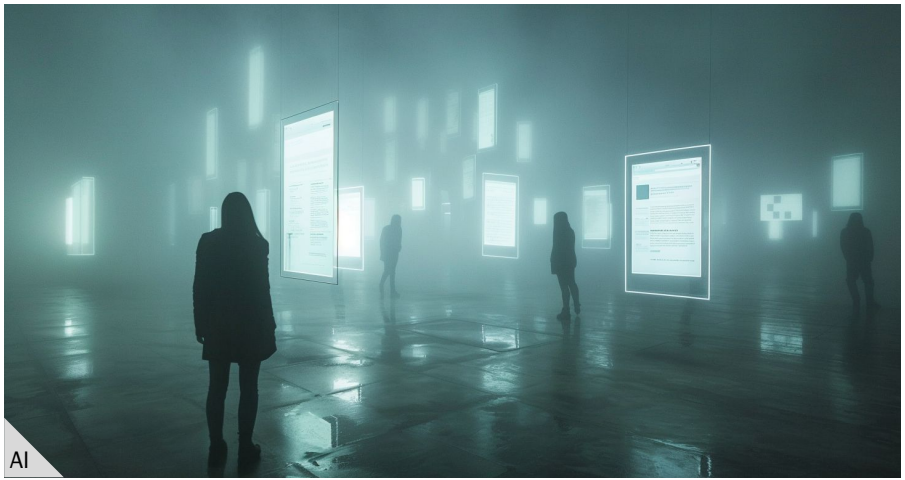


Who's Behind This Website?

https://bit.ly/nicar24-behind-this-website

Priyanjana Bengani
Senior Research Fellow
Tow Center for Digital Journalism
@acookiecrumbles@indieweb.social

Jon Keegan
Investigative Data Journalist
The Markup
@jonkeegan@mastodon.social

Updated March 2024

Agenda

- Asking the right questions: Who? Why? When? How?
- Where to begin?
  - Site content
  - Infrastructure & Networking
  - Connections with Platforms
- Case studies - (Content warning)
- Tools & resources

The new normal: more opacity, less transparency

Existing tools are breaking

- Google Cache no longer exists
- Google dorks / search operators are less reliable
- Dnslytics' tools have moved to "legacy"

Concealing identities is easier

- WHOIS is less useful for new domains due to GDPR
- Squarespace, Wordpress, GoDaddy static websites share IP addresses with thousands of websites
- CDNs make IP address matching harder
- Relatively easy to incorporate a company incognito or with false identities

World of platforms and walled gardens

- Google Analytics 4 disallows finding relationships between websites based on the ID
- Platforms are restricting their transparency tools (TikTok, Twitter, the imminent death of CrowdTangle)
- Far more platforms (Mastodon, BlueSky, Threads, Discord, ...) to monitor
- Facebook's "Related Pages" no longer work

Somebody set this website up!

Who? Why? When? How?

Who?

Who's featured on the website?

Are there authors? Email addresses? Profile pictures?

Are there payment options (crypto, PayPal, donations, subscriptions)? Who's receiving the money?

Are authors common across multiple sites or exclusive to this one?

Is the owner trying to stay hidden?

Why?

Why was this website set up?

To make money through scams or ads (content farms)?

To perform some kind of influence operations?

To promote political candidates or social advocacy?

To deceive the audience by impersonating another website?

When?

When was the domain first registered?

How long has the site existed in its current form?

Was it offline for any significant period of time?

Did the ownership change? (Check historical WHOIS)

Did the site's design or content change drastically at any point?

How?

What is the tech stack?

Where is it hosted?

Wordpress? (Check authors, templates, plugins)

How is the site monetized? Affiliate links, advertising?

Is the content generated by AI?

Where does it link to, and who links to it?

Documenting, archiving, monitoring

YOU NEED TO BE ABLE TO RETRACE YOUR STEPS AND HAVE ALL THE RECEIPTS
ESPECIALLY IN AN ENVIRONMENT WHERE EVERYTHING CHANGES ALL THE TIME

📝 Documenting

- Maintain a **data diary** with detailed notes about what you've looked at and how you got there
- Try to **create a timeline** of the website and how it's evolving over the course of your investigation
- Use Hunchly or screen recordings to keep track of everything you're doing

🕵️ Monitoring

- Set up Klaxon Cloud or VisualPing to be notified of any changes to a site
- Use GitHub Actions and ShotScraper for automated screenshots over time

📚 Archiving

- Archive sites consistently, and in some cases, use multiple archival services (archive.org, archive.is)
- For **public records** or **social media posts**, take screenshots — some of them might not be archivable
- **Download videos** lest they get taken down (Youtube: yt-dlp)
- Take **screenshots with timestamps** so you can monitor changes and gather receipts (GoFullPage).
- Capturing the **full browser window with the URL** field helps strengthen your evidence

So many tools

A word about investigative tools…

TECHNIQUES > TOOLS

- Spend time learning solid, creative investigative techniques
- Most of the time, you'll have to use multiple tools — even if they purport to do the same thing!
- There are SO many tools, and they come and go
- Tools can be expensive, overpromise, underdeliver — and use dubious means to collect data
- Platforms rise and fall, APIs go away
- Don't get too dependent on one tool

https://start.me/p/7kxyy2/osint-tools-curated-by-lorand-bodo

Protecting yourself online

Each investigation will have a different threat model

If you don't want to **expose your IP address** to the website:

- Use a VPN or Tor (note: some VPNs also track your activity!)
- Use a browser different to your primary browser in incognito / private mode

If you don't want to **expose your email address** to the website:

- Use a different email address when you sign up to newsletters
- Ensure your email client blocks remote content from loading to avoid pixel tracking
- Use the '+' trick to see if your email address is shared with other services

If you need to use a virtual machine, use UTM or VMWare

Is it ethical? Is it legal?

SPEAK TO YOUR EDITORS and LAWYERS!

MAKE SURE YOUR INVESTIGATION IS IN THE PUBLIC INTEREST.

Ethics

- Some tools collect data using shady practices — say buying data from data brokers — and purchasing these tools may help sustain these businesses (...which we probably should be reporting on...)
- At times, identifying who's behind a website might end up in doxxing someone's identity — which is complicated

Legal

- If you're scraping publicly-available content, you're *probably* fine but should still be aware of the Computer Fraud and Abuse Act.

⚠️ Be aware of crossing lines while trying to access data: while publicly-available sites are largely fine, using unauthorized credentials is a legal concern

🙄 But, also, note: the St Louis Post-Dispatch reporter who found SSNs exposed by simply "viewing source" on Missouri's Department of Elementary and Secondary Education's (DESE) website was accused of "hacking"

Site content

Content: Text

Check **text fragments** from articles, “about us” pages, and privacy policies to see if they are unique to the site or duplicated [Use exact string matching on multiple search engines]

Run article text through numerous tools to see if the text is **AI-generated**, but note lots of false positives in these tools [GPTZero, OpenAI’s Text Classifier, ContentScale’s AI Detector, CopyLeaks]

Browse site for any names (including bylines), email addresses, phone numbers, addresses, social media handles, and company names

On finding names, or phone numbers, or e-mail addresses, or...

📲 Use tools like Epieos and haveibeenpwned to **reverse lookup emails and phone numbers**: both will show you other services and platforms on which the email address or phone number might exist. TrueCaller also serves as a **reverse phone book**.

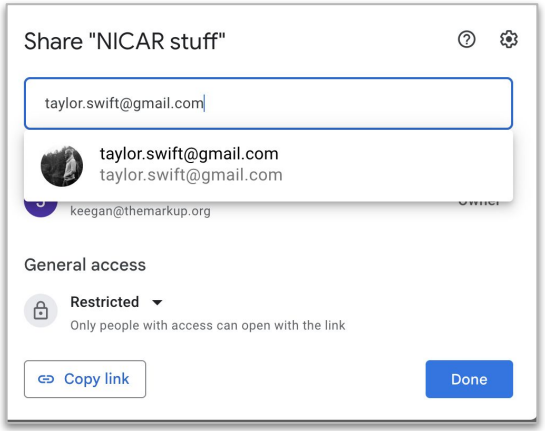
🔎 If you find **social media handles**, use tools like sherlock that will scan multiple platforms to see if the same handle appears elsewhere — you'll still have to confirm that it's the same user and not just the same handle

🪪 If you find names, use Bellingcat's Name Variant Tool to find possible variations on any names

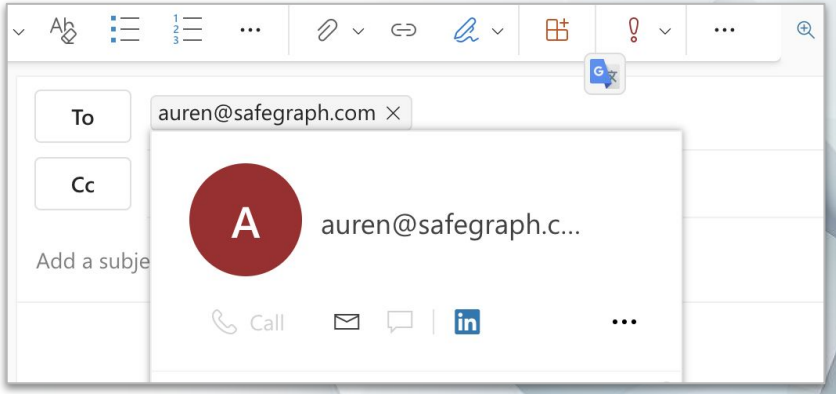
🏢 If you find **company names**, use OpenCorporates or LinkedIn to see whether any personnel information is available. OpenCorporates also lets you **search by addresses** — so you can find who else shares the same office location!

Given an email address...

If you find a Gmail address and the other tools don't give you a break, enter the address in Google Docs' sharing sheet. If the address is valid, you should see a profile picture.
⚠️Remember to not actually share the document!



If you want to find a LinkedIn profile associated with an email address, compose an email in Outlook, and put the address in the "To" field.



Content: Media

🌆 Use **reverse image search** across multiple platforms (Google, Bing, Yandex) to check:

- If image is a stock image
- If the profile picture for an author appears elsewhere
- If the image is posted on other sites

👨 Use **facial recognition tools** (PimEyes, Search4Faces, FaceCheck.ID, Amazon Rekognition) to see if it can identify faces in the profile pictures

📷 Use **EXIF tools** (EXIF Data) to check:

- If any metadata is retained in the images. This can include location, time the photograph was taken, camera details, etc

✂️ Use **forensics tools** to check:

- If media is AI-generated [WeVerify (InVid)]
- If media has been modified from the original [Forensically]

Content: Documents

Google dorking

To find documents that might be hosted on your website, use Google dorking: `filetype:pdf site:<domain_of_interest.com>`. Also try XLS, XLSX, PPT, DOC

If you find a PDF, open it (ideally in a browser) and check document details / information for any names. And check PDF metadata. Use Dangerzone or Google Drive if you are worried about malware.

Amazon S3

Look at the website's source to see if there's anything hosted on AWS S3. The bucket name can be revealing, and the bucket might be open if misconfigured.

Donations, Memberships, Subscriptions

PayPal

💳 If there's a PayPal button, click through until you finally get to the screen that tells you who you're about to pay. If the PayPal account has been banned, you can still try the PayPal link in Internet Archive.

Crypto

💸Lookup the wallet address in search engines to see where else they show up

Lookup the wallet address in **crypto investigation tools** (EtherScan, Breadcrumbs) to see how much money it has and how active it is

Others

💰 Websites might have a plethora of payment services — from gaming platforms to AliPay — and sometimes it's worth going through all the steps to see what information you can find out about the payee. (Case study: Behind a Secretive Global Network of Non-Consensual Deepfake Pornography)

WordPress Websites

Use wpscan to see the **theme** a Wordpress site uses as well as the **authors**

WordPress Themes and plugins are often used across networks of sites. Check Built.with

```
 | Found By: Rss Generator (Passive Detection)
 |  - https://empirenews.net/feed/, <generator>https://wordpress.org/?v=6.4.3</generator>
 |  - https://empirenews.net/comments/feed/, <generator>https://wordpress.org/?v=6.4.3</generator>

[+] WordPress theme in use: empire-news
 | Location: https://empirenews.net/wp-content/themes/empire-news/
 | Readme: https://empirenews.net/wp-content/themes/empire-news/README.md
 | Style URL: https://empirenews.net/wp-content/themes/empire-news/style.css?ver=6.4.3
 | Style Name: empire news
 | Description: Empire News theme...
 | Author: Jitender Thakur
 |
 | Found By: Css Style In Homepage (Passive Detection)
 | Confirmed By: Css Style In 404 Page (Passive Detection)
 |
 | Version: 1.0 (80% confidence)
 | Found By: Style (Passive Detection)
 |  - https://empirenews.net/wp-content/themes/empire-news/style.css?ver=6.4.3, Match: 'Version: 1.0'
```

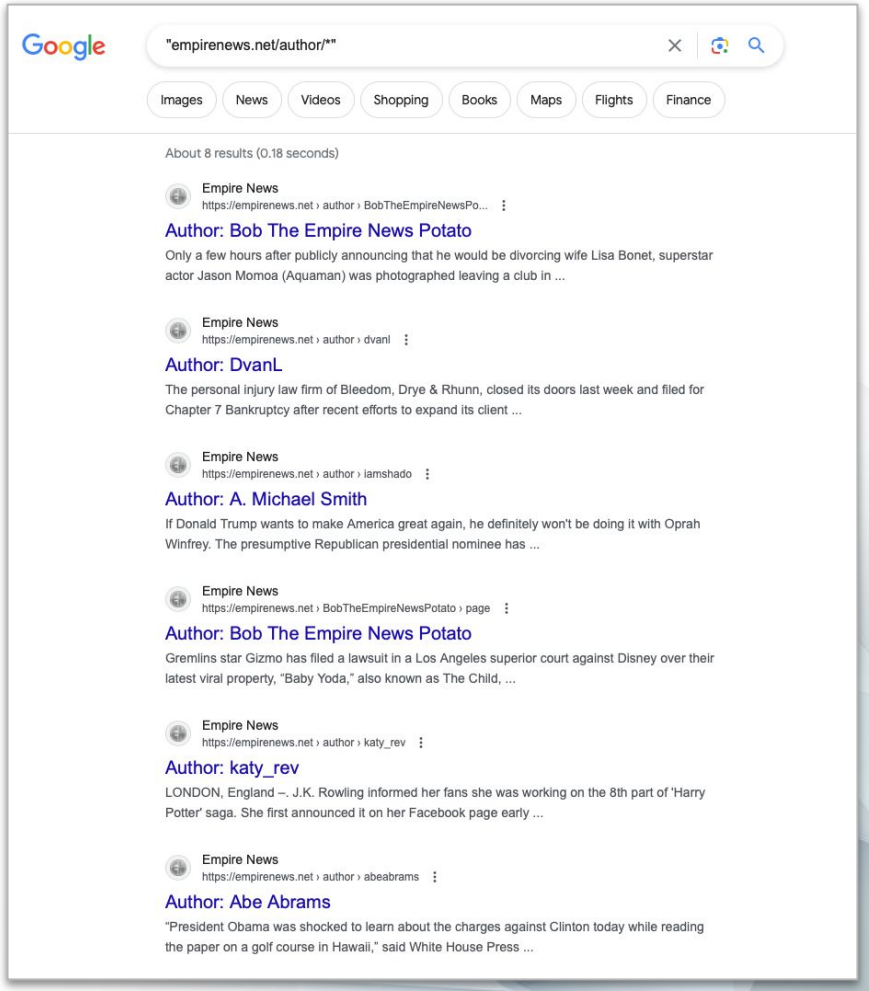
https://wpscan.com/

WordPress Websites

Use the Google author trick:
"`empirenews.net/author/*`"

Pulling RSS feeds can sometimes show you authors

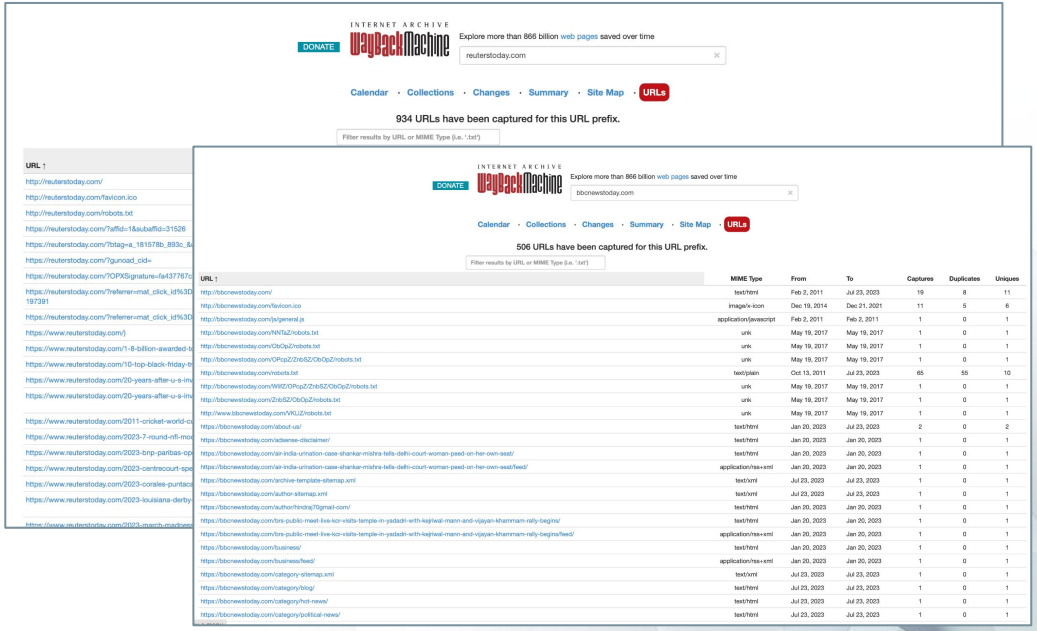
You can pull RSS article links into Google Sheets using `IMPORTFEED`



Internet Archive

To get the lay of the land of a website, look at the URLs tab of the website, which shows all the URLs Internet Archive has captured, along with dates and number of captures.

This waybackurls.py script will do this for you quickly.



https://web.archive.org/web

Reuters Today and BBC News Today were identified by Bleeping Computer as part of a content farm impersonating major news outlets

Infrastructure and networking

When it comes to domains, we need to ask

🧑‍💼 Ownership

Who owns the domain?

What other domains do they own?

What network is it a part of, if any?

🌐 Links and network requests

Which domains do they link to?

Which domains link to it?

Where do images and other media load from?

💥 Shared traits

Which other domains share the IP addresses?

Which domains share tracking & adtech identifiers?

Which domains share SSL certs or favicons?

💰 Adtech

Which countries' ad-tech platforms are used?

Who are the top advertisers / advertising platforms?

Who owns the domain?
What other domains do they own?

Lookup the **WHOIS data** for the domain (current and historic)

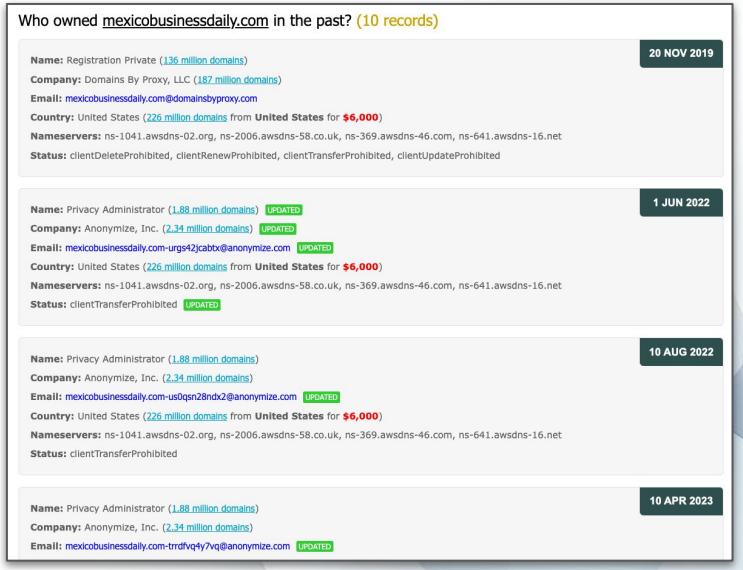
GDPR has made this useless for new domains, but Whoxy provides historic domain registration details

Even if registration details are obscured, you still see **when a domain was registered** and with **whom**

People can lie in WHOIS — it's up to the domain registrar to check

To find other possible related domains, use dnstwist, which gives you domain permutations

Tools: Whoxy, RiskIQ, DomainTools, DNSTwist



https://www.whoxy.com/

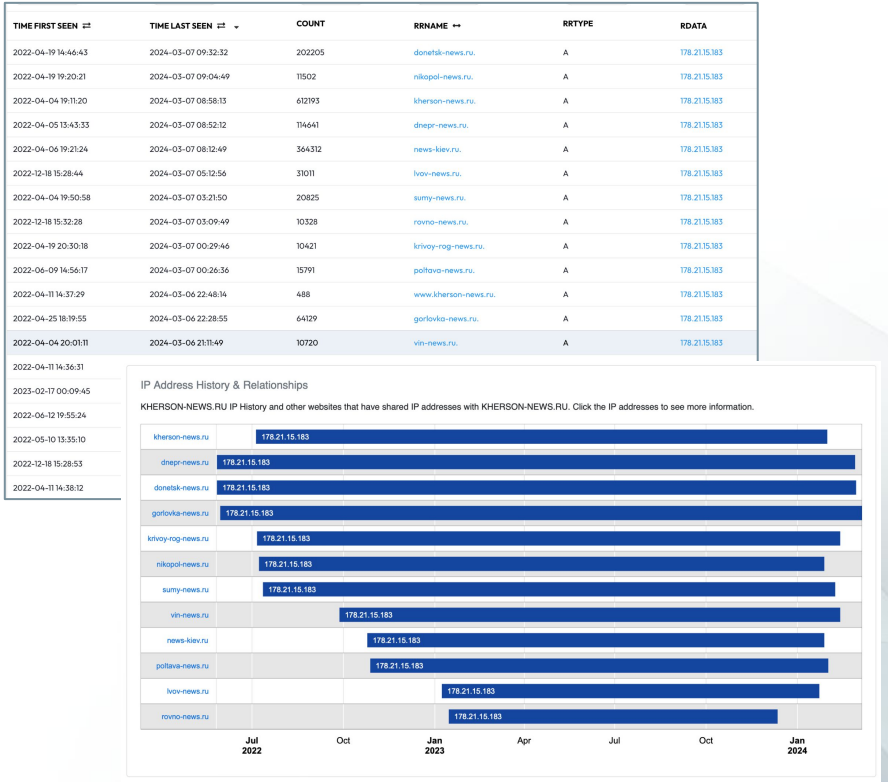
Which other domains share the IP addresses?

Find the current and historic IP addresses of the domain.

Find what other domains reside — or have resided — on the same IP address over what time period

If the IP address is Cloudflare or Google Cloud or GoDaddy, matching IP addresses aren't a reliable indicator — if you get thousands of results, abort!

Tools: Built.With, SecurityTrails, RiskIQ, DomainTools, FarSight DNSDB



TIME FIRST SEEN	TIME LAST SEEN	COUNT	RRNAME	RRTYPE	RDATA
2022-04-19 14:46:43	2024-03-07 09:32:32	202205	donetsk-news.ru.	A	178.21.15.183
2022-04-19 19:20:21	2024-03-07 09:04:49	11502	nikopol-news.ru.	A	178.21.15.183
2022-04-04 19:11:20	2024-03-07 08:58:13	612193	kherson-news.ru.	A	178.21.15.183
2022-04-05 13:43:33	2024-03-07 08:52:12	114641	dnepr-news.ru.	A	178.21.15.183
2022-04-06 19:21:24	2024-03-07 08:12:49	364312	news-kiev.ru.	A	178.21.15.183
2022-12-18 15:28:44	2024-03-07 05:12:56	31011	lvov-news.ru.	A	178.21.15.183
2022-04-04 19:50:58	2024-03-07 03:21:50	20825	sumy-news.ru.	A	178.21.15.183
2022-12-18 15:32:28	2024-03-07 03:09:49	10328	rovno-news.ru.	A	178.21.15.183
2022-04-19 20:30:18	2024-03-07 00:29:46	10421	krivoy-rog-news.ru.	A	178.21.15.183
2022-06-09 14:56:17	2024-03-07 00:26:36	15791	poltava-news.ru.	A	178.21.15.183
2022-04-11 14:37:29	2024-03-06 22:48:14	488	www.kherson-news.ru.	A	178.21.15.183
2022-04-25 18:19:55	2024-03-06 22:28:55	64129	gorlovka-news.ru.	A	178.21.15.183
2022-04-04 20:01:11	2024-03-06 21:11:49	10720	vin-news.ru.	A	178.21.15.183
2022-04-11 14:36:31					
2023-02-17 00:09:45					
2022-06-12 19:55:24					
2022-05-10 13:35:10					
2022-12-18 15:28:53					
2022-04-11 14:38:12					

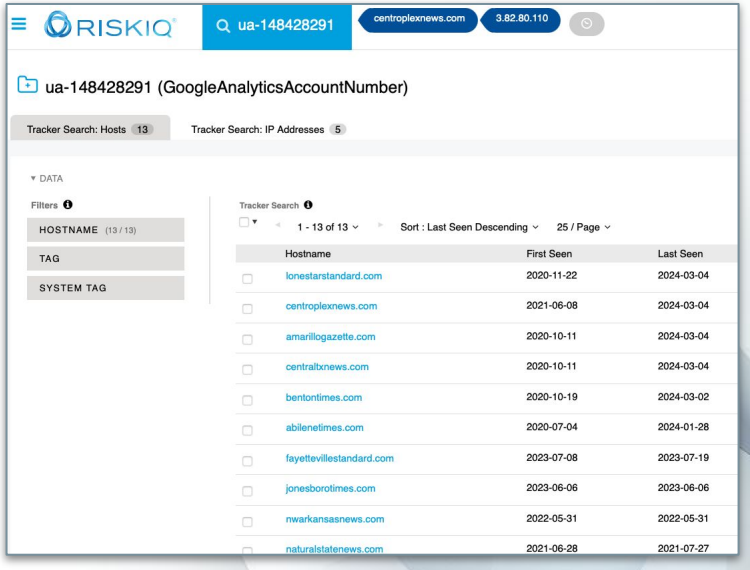
https://built.with & DNSDB Scout

Which other domains share the same analytics identifiers?

Identify third-party analytics and tracking libraries used by the website. These include Google Analytics, Facebook Pixel, Quantcast, NewRelic, Google Tag Manager, etc.

Find other domains that share the same identifiers

Tools: RiskIQ, BuiltWith



https://community.riskiq.com

A note about Google Analytics

Google Analytics IDs used to be in the syntax **`UA-NUMBER-SUFFIX`**. Related domains shared the same **`UA-NUMBER`**, but the SUFFIX was different.

Google Analytics 4 changes this such that each website has different identifiers

Bellingcat's Wayback Analytics tool looks at Internet Archive to extract old analytics IDs, which might be helpful in the present and can be used to aid the investigation

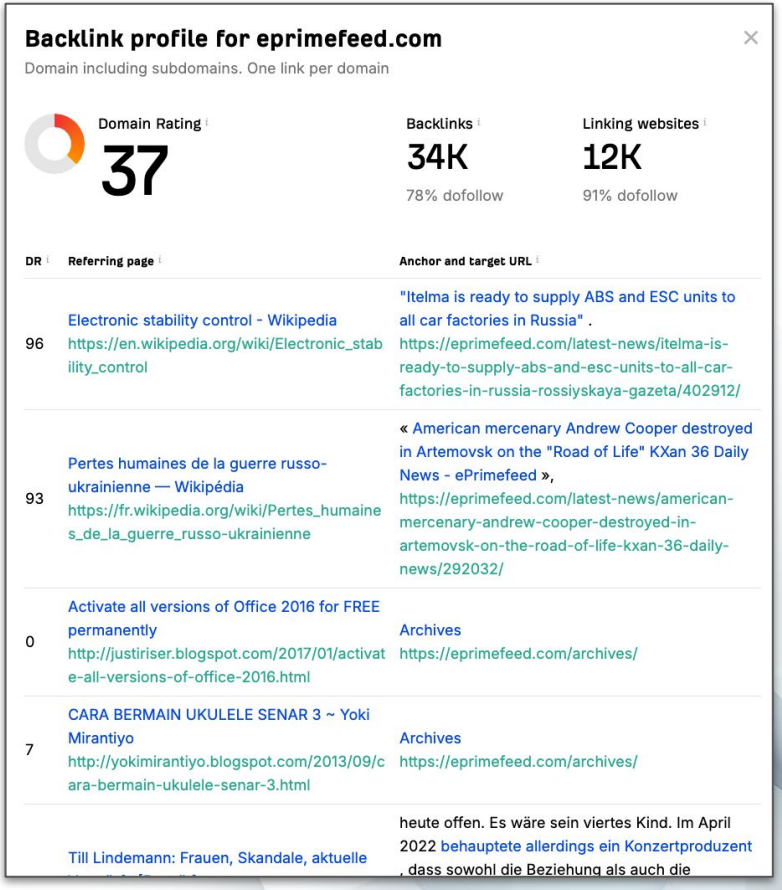
https://github.com/bellingcat/wayback-google-analytics

Links

Find who's linking to the website of interest consistently by using a **backlink checker** (ahrefs, Moz) — what's the relationship between the sites?

Use photon or urlscan.io to gather the **outbound urls**, (urls a site links to), as well as some high-level "intel" — who's the site linking to the most?

Analyze outbound links, especially those to merch stores, for **affiliate links** — who's the affiliate? (Especially useful for health and wellness scams)



https://ahrefs.com/backlink-checker

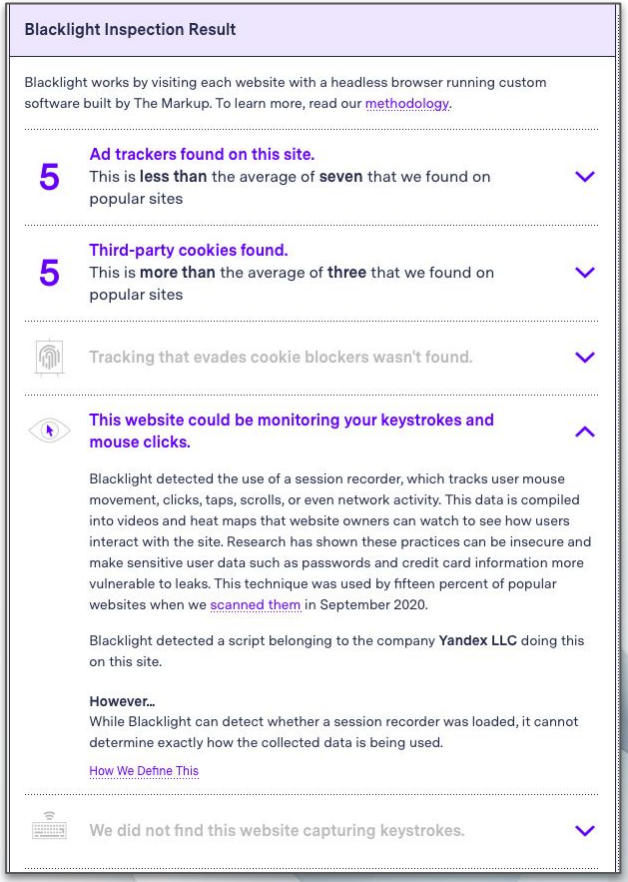
Tracking, Analytics, AdTech

Analyzing the stack

Which ad trackers, third-party cookies, and other creepy surveillance libraries is the site using?

This might provide some insight into the motivations behind the site.

Tools: Blacklight, BuiltWith



https://themarkup.org/blacklight

Tracking, Analytics, AdTech

Monitoring network requests

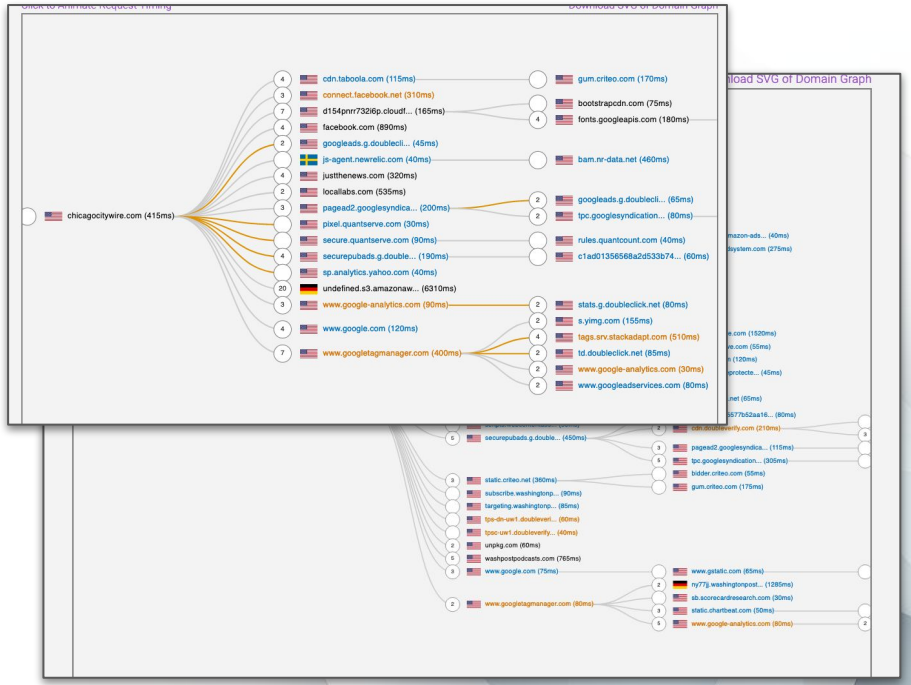
Is the site making network calls to ad-tech providers in other countries?

Is the url for its static files (images, videos) different to the website?

Are there any network calls that warrant further investigation? (Check the Network tab in developer tools)

Who are the top advertisers?

Use [FouAnalytics X-Ray] to see all the network requests a site makes.



https://pagexray.fouanalytics.com/

Connections to platforms

When it comes to platforms, we need to ask

TOO MANY PLATFORMS, TOO LITTLE TIME

🚉 Individual platform

- On which platforms does the domain get high traction? And is it on the domain's account or elsewhere?
- Which other accounts does it interact with?
- What other domains does it promote?
- Does it look like engagement was purchased?
- Where else does the profile picture appear?
- Does the account appear to have been co-opted at any point?

💥 Presence on other platforms

- Even if not listed on the site, is there presence on other platforms?
- Tools like Sherlock and Blackbird help you find other social media accounts

💰 Ads

- Are there ads featuring the domain on any platforms?
- Which entity has run those ads?

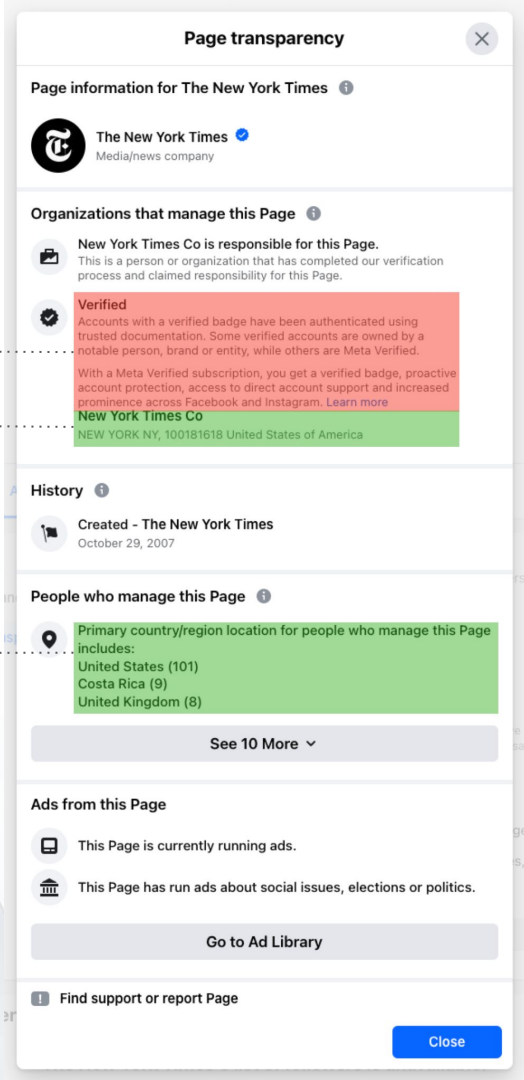
Facebook Page Transparency

On any Facebook Page, go to “About” and then “Page Transparency,” and then click on “See All”

- largely pointless as it could be a notable entity or it could be someone who paid Meta
- addresses can be plugged into OpenCorporates to see which other entities share the address
- useful to understand where page managers are located, and if that aligns with what the page says it does

These fields are typically only filled in for companies or pages running ads

https://facebook.com/

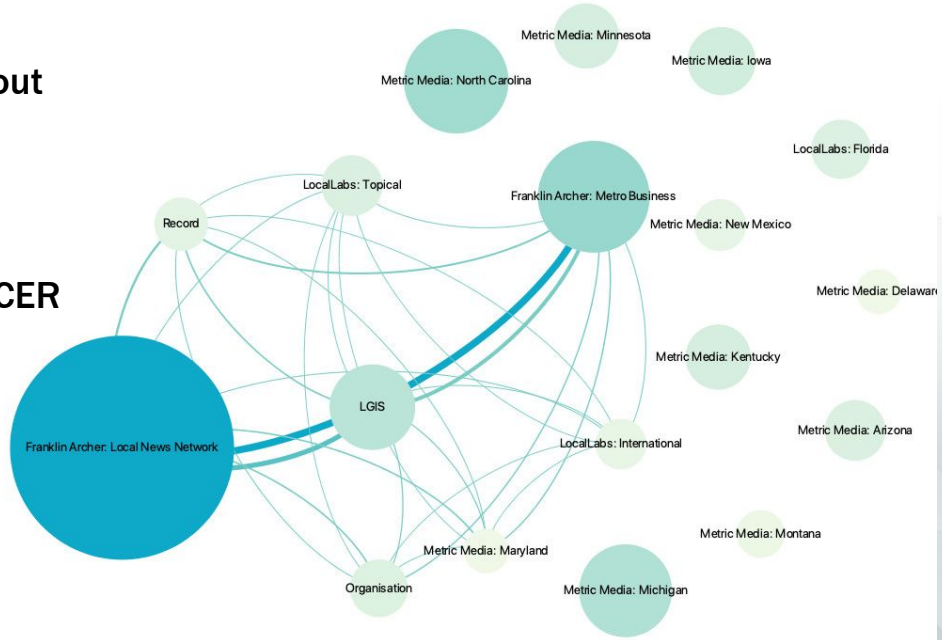


Case studies

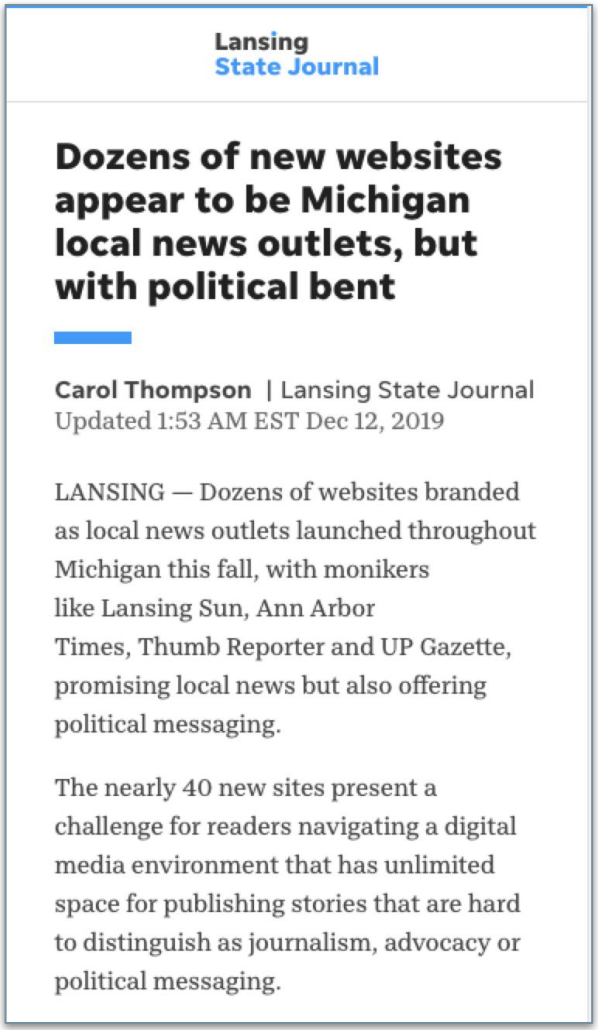
https://aitestkitchen.withgoogle.com/tools/image-fx

Investigating politically-backed “pinkslime” local news

- Analyzing domains: **RiskIQ, FarSight DNSDB Scout**
- Analyzing matching analytics: **RiskIQ**
- Following the money: **OpenSecrets, FEC filings, state records, 990s**
- Corporate ownership web: **OpenCorporates, PACER**
- Collaborations & partners: **Facebook and Google Ad Library, DNSDB Flexible Search**



https://www.cjr.org/tow_center_reports/hundreds-of-pink-slime-local-news-outlets-are-distributing-algorithmic-stories-conservative-talking-points.php



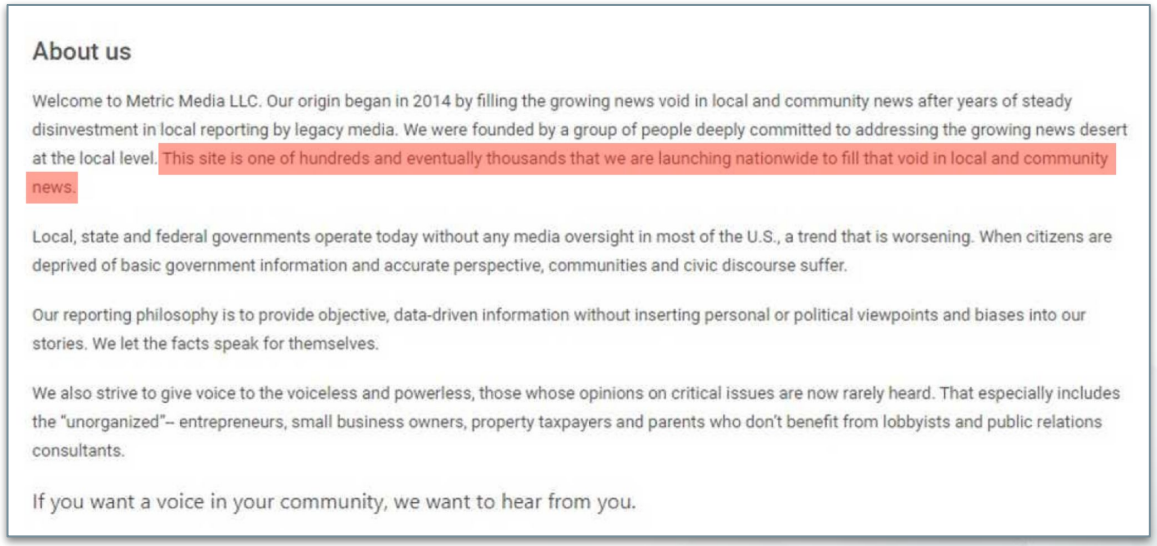
Lansing
State Journal

Dozens of new websites appear to be Michigan local news outlets, but with political bent

Carol Thompson | Lansing State Journal
Updated 1:53 AM EST Dec 12, 2019

LANSING — Dozens of websites branded as local news outlets launched throughout Michigan this fall, with monikers like Lansing Sun, Ann Arbor Times, Thumb Reporter and UP Gazette, promising local news but also offering political messaging.

The nearly 40 new sites present a challenge for readers navigating a digital media environment that has unlimited space for publishing stories that are hard to distinguish as journalism, advocacy or political messaging.



About us

Welcome to Metric Media LLC. Our origin began in 2014 by filling the growing news void in local and community news after years of steady disinvestment in local reporting by legacy media. We were founded by a group of people deeply committed to addressing the growing news desert at the local level. This site is one of hundreds and eventually thousands that we are launching nationwide to fill that void in local and community news.

Local, state and federal governments operate today without any media oversight in most of the U.S., a trend that is worsening. When citizens are deprived of basic government information and accurate perspective, communities and civic discourse suffer.

Our reporting philosophy is to provide objective, data-driven information without inserting personal or political viewpoints and biases into our stories. We let the facts speak for themselves.

We also strive to give voice to the voiceless and powerless, those whose opinions on critical issues are now rarely heard. That especially includes the "unorganized"-- entrepreneurs, small business owners, property taxpayers and parents who don't benefit from lobbyists and public relations consultants.

If you want a voice in your community, we want to hear from you.

🤔 How do you find the thousands of political outlets launching to fill the void in local and community news?

Looking up the IP addresses of a couple of domains in **DNSDB Scout** yielded over a thousand results!

Show 25 entries

To Unicode | Invert | Export

Filter Time First	Filter Time Last	Filter Count	Filter RRName	Filter RRType	Filter RData
Time First Seen	Time Last Seen	Count	RRName	RRType	RData
2020-01-18 18:21:08	2020-09-06 03:06:46	199	youngstowntimes.com.	A	3.218.216.245
2020-01-18 19:34:12	2020-09-03 18:08:43	203	wcindiananews.com.	A	3.218.216.245
2020-01-18 18:16:41	2020-09-03 13:13:30	217	whitewatertimes.com.	A	3.218.216.245
2020-01-18 18:16:48	2020-09-03 07:32:56	172	weststarknews.com.	A	3.218.216.245
2020-01-18 18:16:43	2020-09-03 07:30:28	228	westlucasnews.com.	A	3.218.216.245
2020-01-18 19:34:21	2020-09-03 07:30:10	190	westindynews.com.	A	3.218.216.245
2020-01-18 18:17:47	2020-09-03 07:30:02	215	westhamiltonnews.com.	A	3.218.216.245
2019-09-15 13:13:30	2020-09-03 07:29:49	582	westernwaynetoday.com.	A	3.218.216.245
2020-01-21 18:25:20	2020-09-03 07:29:05	211	westclevelandnews.com.	A	3.218.216.245
2020-01-18 18:16:47	2020-09-03 06:18:56	207	wcwisconsinnews.com.	A	3.218.216.245
2019-09-08 20:00:29	2020-09-03 06:18:40	686	wcmichigannews.com.	A	3.218.216.245
2020-01-18 18:24:07	2020-09-03 06:08:23	269	waynecountytoday.com.	A	3.218.216.245
2020-01-18 18:16:41	2020-09-03 06:06:31	202	waukeshatimes.com.	A	3.218.216.245
2019-09-08 20:00:36	2020-09-03 06:05:03	562	waterfordtoday.com.	A	3.218.216.245
2020-01-18 18:16:46	2020-09-03 05:55:09	213	washingtoncotimes.com.	A	3.218.216.245
2019-09-05 20:25:08	2020-09-03 05:53:47	780	warrensun.com.	A	3.218.216.245
2020-01-18 18:16:45	2020-09-03 05:53:39	223	warrenclintonnews.com.	A	3.218.216.245
2019-09-05 21:14:25	2020-09-03 02:33:22	754	thumbreporter.com.	A	3.218.216.245
2020-01-18 18:16:47	2020-09-03 01:46:18	186	vacationlandtimes.com.	A	3.218.216.245
2019-09-05 21:14:44	2020-09-03 00:56:20	758	upgazette.com.	A	3.218.216.245
2020-01-18 18:16:42	2020-09-02 23:00:14	187	tuscarawasnews.com.	A	3.218.216.245
2020-01-21 05:55:15	2020-09-02 22:17:35	227	trumballnews.com.	A	3.218.216.245
2020-01-18 18:20:32	2020-09-02 21:45:18	252	triconews.com.	A	3.218.216.245
2019-09-08 19:32:33	2020-09-02 21:45:05	617	tricitysun.com.	A	3.218.216.245
2020-01-18 18:32:27	2020-09-02 19:58:40	206	toledoreporter.com.	A	3.218.216.245

1 to 25 of 137 Results

First Previous 1 2 3 4 5 6 Next Last

Show 25 entries

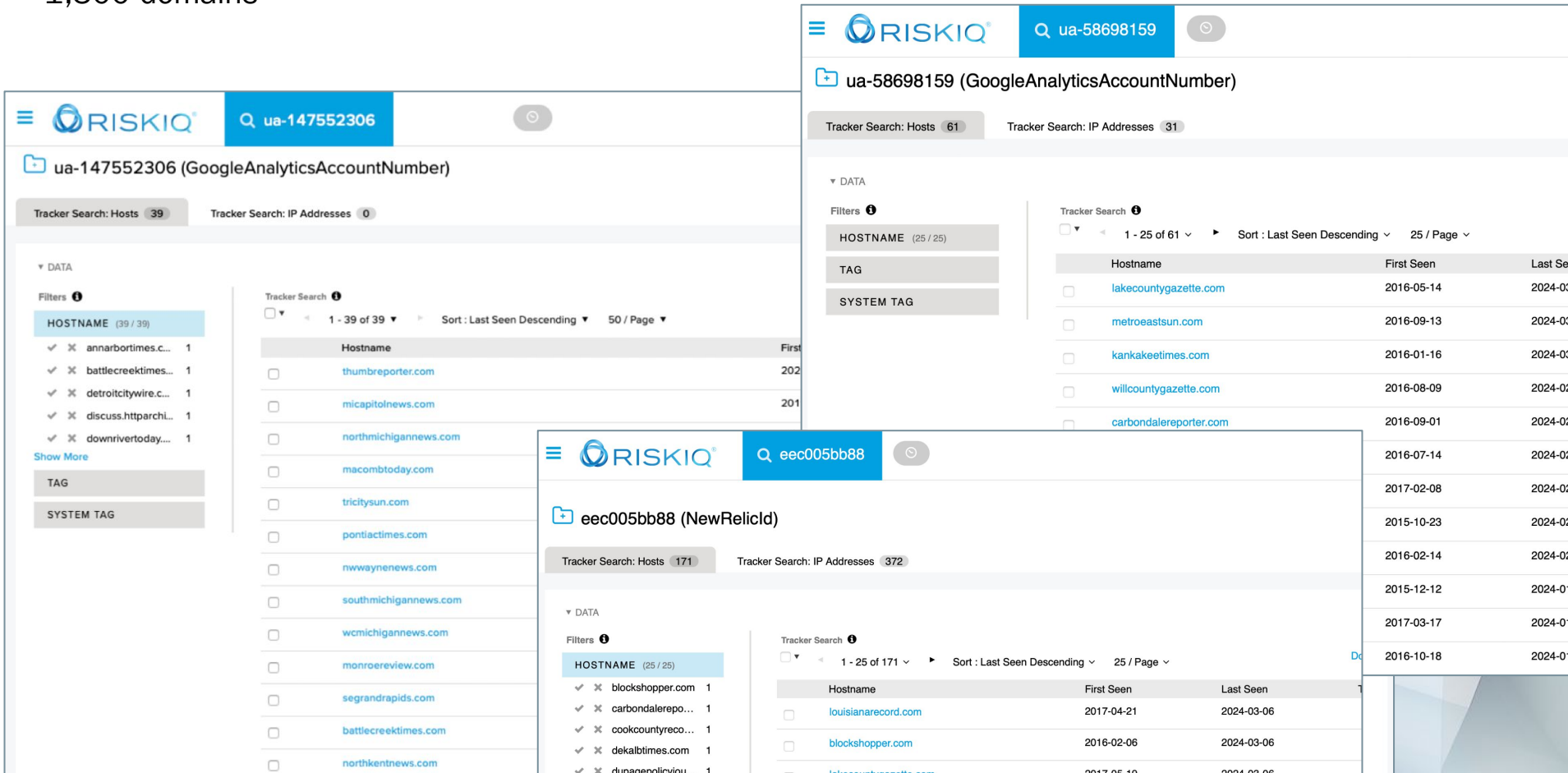
To Unicode | Invert | Export

Filter Time First	Filter Time Last	Filter Count	Filter RRName	Filter RRType	Filter RData
Time First Seen	Time Last Seen	Count	RRName	RRType	RData
2020-09-05 04:28:11	2022-03-03 23:33:53	118	southernwvnews.com.	A	34.236.176.60
2020-08-31 08:51:16	2022-03-03 23:29:07	591	ftworthtimes.com.	A	34.236.176.60
2020-08-28 23:26:10	2022-03-03 23:27:44	539	easthoustonnews.com.	A	34.236.176.60
2020-08-27 22:00:03	2022-03-03 23:25:35	5319	centralutahnews.com.	A	34.236.176.60
2020-09-09 06:08:33	2022-03-03 23:19:43	5815	nefranklinnews.com.	A	34.236.176.60
2020-09-13 19:01:03	2022-03-03 23:08:57	16685	southdaytonnews.com.	A	34.236.176.60
2020-08-30 01:37:56	2022-03-03 22:57:45	6275	lowedeltanews.com.	A	34.236.176.60
2020-08-24 15:07:46	2022-03-03 22:55:38	3358	jonesborotimes.com.	A	34.236.176.60
2020-09-14 03:50:39	2022-03-03 22:55:00	3633	larimernews.com.	A	34.236.176.60
2020-08-26 01:37:43	2022-03-03 22:54:52	7815	eastclevelandnews.com.	A	34.236.176.60
2020-08-21 09:38:04	2022-03-03 22:51:33	6183	beehivestatenews.com.	A	34.236.176.60
2020-09-03 18:08:43	2022-03-03 22:51:14	4171	southutahnews.com.	A	34.236.176.60
2020-09-04 09:56:58	2022-03-03 22:49:21	159	northokcnews.com.	A	34.236.176.60
2020-08-24 19:10:08	2022-03-03 22:45:16	493	coachellatoday.com.	A	34.236.176.60
2020-09-16 05:44:25	2022-03-03 22:41:54	128	westoctimes.com.	A	34.236.176.60
2020-09-10 06:08:28	2022-03-03 22:40:56	393	tulsastandard.com.	A	34.236.176.60
2020-09-16 10:44:44	2022-03-03 22:33:19	9300	sekentuckynews.com.	A	34.236.176.60
2020-09-11 17:26:10	2022-03-03 22:32:33	160	phillyleader.com.	A	34.236.176.60
2020-08-21 09:06:35	2022-03-03 22:25:30	3171	annarbortimes.com.	A	34.236.176.60
2020-09-16 23:22:50	2022-03-03 22:23:08	15366	swgeorgianews.com.	A	34.236.176.60
2020-09-14 17:14:11	2022-03-03 22:21:56	7617	nesacramentonews.com.	A	34.236.176.60
2020-09-15 04:53:10	2022-03-03 22:21:28	6005	nepiedmontnews.com.	A	34.236.176.60
2020-09-12 13:09:50	2022-03-03 22:18:44	3870	georgiamountainnews.com.	A	34.236.176.60
2020-08-24 18:26:12	2022-03-03 22:18:23	4922	ibxnews.com.	A	34.236.176.60
2020-09-07 04:06:48	2022-03-03 22:14:41	1708	southpimanews.com.	A	34.236.176.60

1 to 25 of 960 Results

First Previous 1 2 3 4 5 6 7 ... 39 Next Last

Putting each domain through RiskIQ to find tracking identifiers (Google, Facebook, NewRelic), we identified ~1,300 domains



PIPELINE MEDIA LLC

Company Number	LLC_02632055
Native Company Number	02632055
Status	Goodstanding
Incorporation Date	10 November 2008 (over 15 years ago)
Company Type	Limited Liability Company
Jurisdiction	Illinois (US)
Registered Address	2045 W. GRAND AVE. STE B CHICAGO 60612 IL United States
Previous Names	LOCALITY LABS, LLC JOURNATIC, LLC BLOCKSHOPPER LLC
Alternative Names	PIPELINE MEDIA SERVICES LLC (trading name, 2021-07-15 -)
Agent Name	L. ROBERT PASQUESI
Agent Address	585 BANK LN. STE 3000, LAKE FOREST, Sangamon, 60045
Directors / Officers	CAMERON, BRAD, manager DUNN, TIMOTHY, manager L. ROBERT PASQUESI, agent TIMPONE, BRIAN, manager
Registry Page	https://apps.ilsos.gov/corporatellc/ *Website unavailable outside U.S.*

Events for PIPELINE MEDIA LLC

On 2008-11-10	Incorporated
Between 2008-11-10 and 2022-11-07	Addition of officer CAMERON, BRAD, manager
Between 2008-11-10 and 2022-11-07	Addition of officer DUNN, TIMOTHY, manager
Between 2008-11-10 and 2022-11-07	Addition of officer L. ROBERT PASQUESI, agent
Between 2008-11-10 and 2022-11-07	Addition of officer TIMPONE, BRIAN, manager

🤔 How do you confirm this is actually the West Texas oil billionaire who's heavily involved in Texas politics and he already has existing close ties to the network?

⏰ When did Dunn get involved with this network?

https://opencorporates.com

ANALYSIS

As election looms, a network of mysterious 'pink slime' local news outlets nearly triples in size

AUGUST 4, 2020
By PRIYANJANA BENGANI

TOW REPORT

Hundreds of 'pink slime' local news outlets are distributing algorithmic stories and conservative talking points

DECEMB
By PRIYA

TOW REPORT

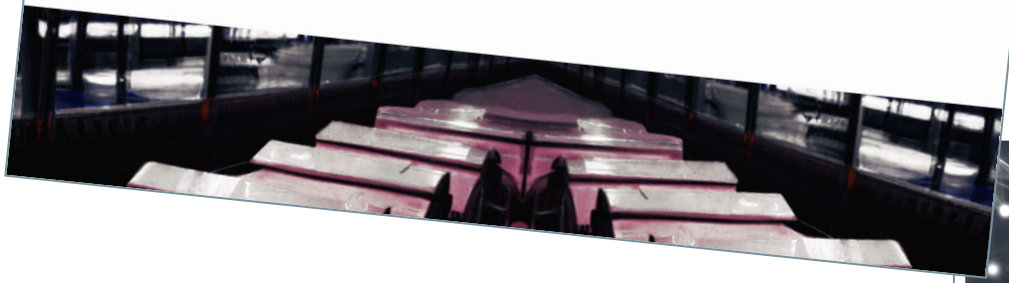
Advocacy groups and Metric Media collaborate on local 'community news'

OCTOBER 14, 2021
By PRIYANJANA BENGANI

CJR PARTNER | TOW CENTER

'Pink slime' network gets $1.6M election boost from PACs backed by oil-and-gas, shipping magnates

OCTOBER 31, 2022
By PRIYANJANA BENGANI



TOW REPORT

The Metric Media network runs more than 1,200 local news sites. Here are some of the non-profits funding them.

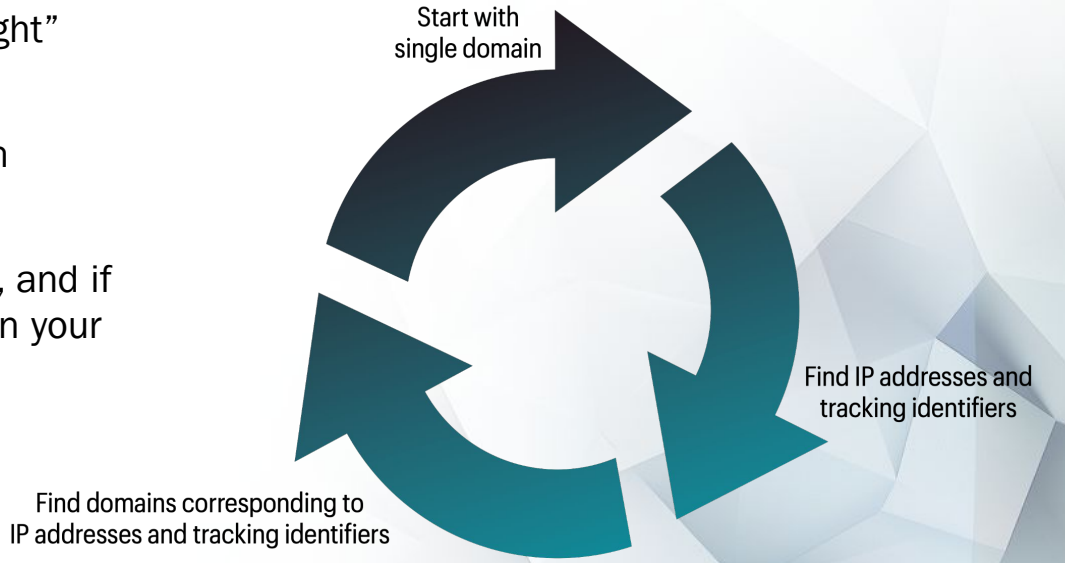
OCTOBER 14, 2021
By PRIYANJANA BENGANI

Everything that goes into the story has to be verified — maybe manually

Just because you identified a person's name doesn't necessarily mean you found the "right" person

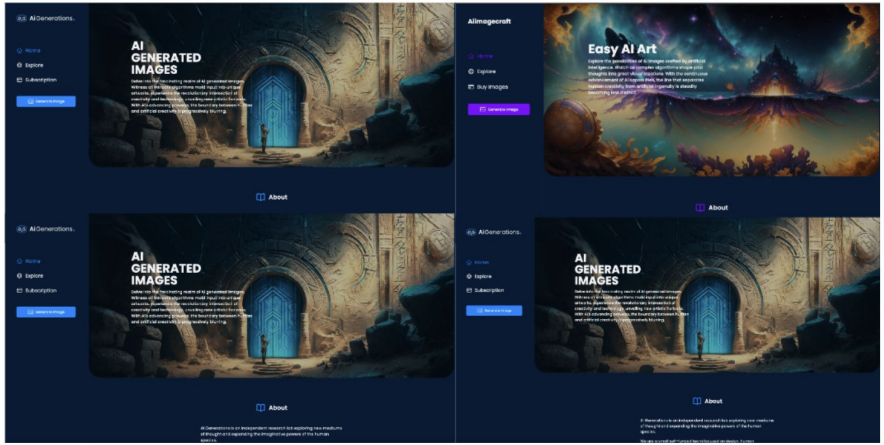
Google obsessively — you never know when something new comes up

Document everything, things change rapidly, and if you don't have the receipts, you can't rely on your memory!



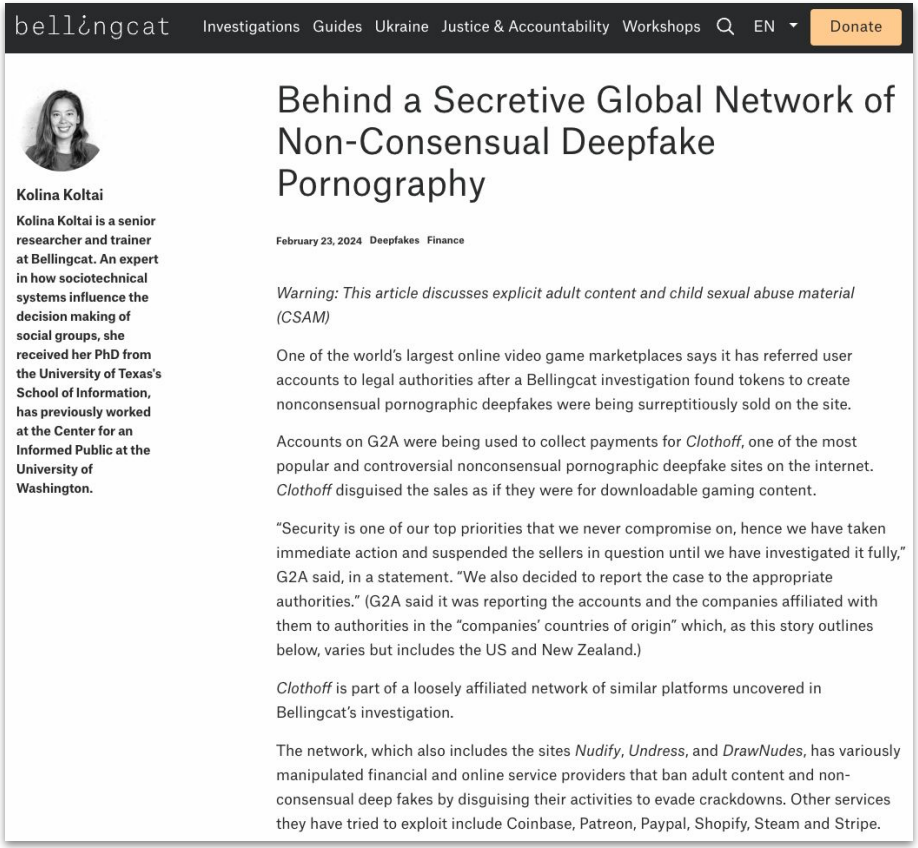
Bellingcat: Pornographic Deepfake Sites

By Kolina Koltai



https://www.bellingcat.com

- Analyzing domains: **DomainTools**
- Analyzing photographs
- Following the money across payment and e-commerce platforms: **Steam, G2A, PayPal, Shopify, Coinbase, Patreon, Venmo, Stripe**
- Revealing hidden information from archives
- Investigating business records in several regions
- Connecting social media profiles: **Facebook, LinkedIn, VKontakte**
- Identifying AI generated images and text
- Following network redirects



bellingcat Investigations Guides Ukraine Justice & Accountability Workshops EN Donate

Behind a Secretive Global Network of Non-Consensual Deepfake Pornography

Kolina Koltai

Kolina Koltai is a senior researcher and trainer at Bellingcat. An expert in how sociotechnical systems influence the decision making of social groups, she received her PhD from the University of Texas's School of Information, has previously worked at the Center for an Informed Public at the University of Washington.

February 23, 2024 Deepfakes Finance

Warning: This article discusses explicit adult content and child sexual abuse material (CSAM)

One of the world's largest online video game marketplaces says it has referred user accounts to legal authorities after a Bellingcat investigation found tokens to create nonconsensual pornographic deepfakes were being surreptitiously sold on the site.

Accounts on G2A were being used to collect payments for *Clothoff*, one of the most popular and controversial nonconsensual pornographic deepfake sites on the internet. *Clothoff* disguised the sales as if they were for downloadable gaming content.

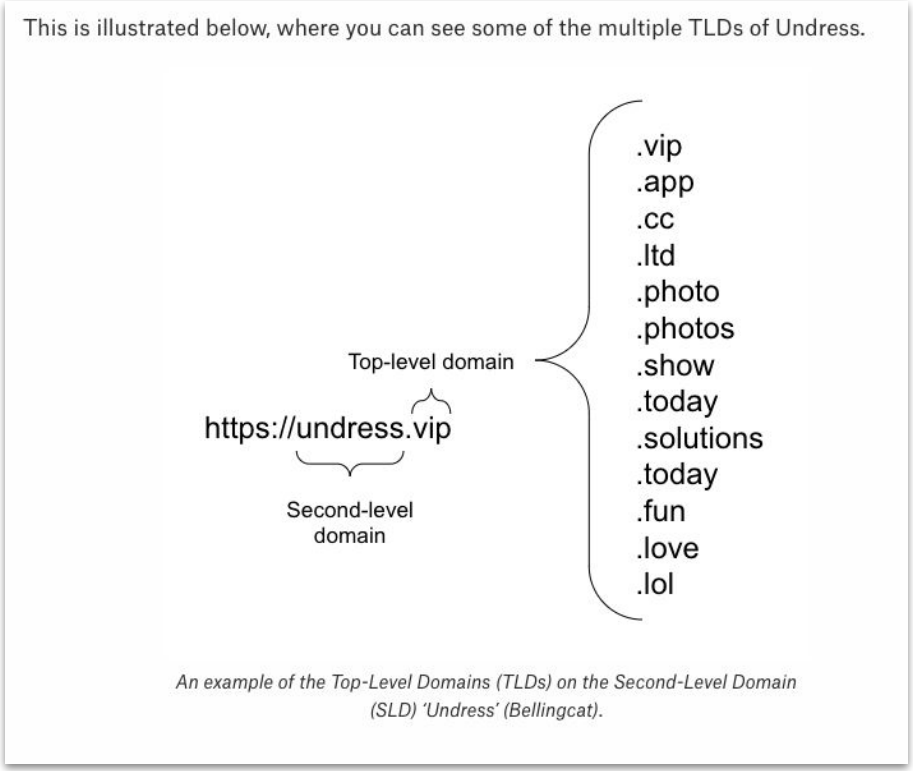
"Security is one of our top priorities that we never compromise on, hence we have taken immediate action and suspended the sellers in question until we have investigated it fully," G2A said, in a statement. "We also decided to report the case to the appropriate authorities." (G2A said it was reporting the accounts and the companies affiliated with them to authorities in the "companies' countries of origin" which, as this story outlines below, varies but includes the US and New Zealand.)

Clothoff is part of a loosely affiliated network of similar platforms uncovered in Bellingcat's investigation.

The network, which also includes the sites *Nudify*, *Undress*, and *DrawNudes*, has variously manipulated financial and online service providers that ban adult content and non-consensual deep fakes by disguising their activities to evade crackdowns. Other services they have tried to exploit include Coinbase, Patreon, Paypal, Shopify, Steam and Stripe.

https://www.bellingcat.com/news/2024/02/23/behind-a-secretive-global-network-of-non-consensual-deepfake-pornography

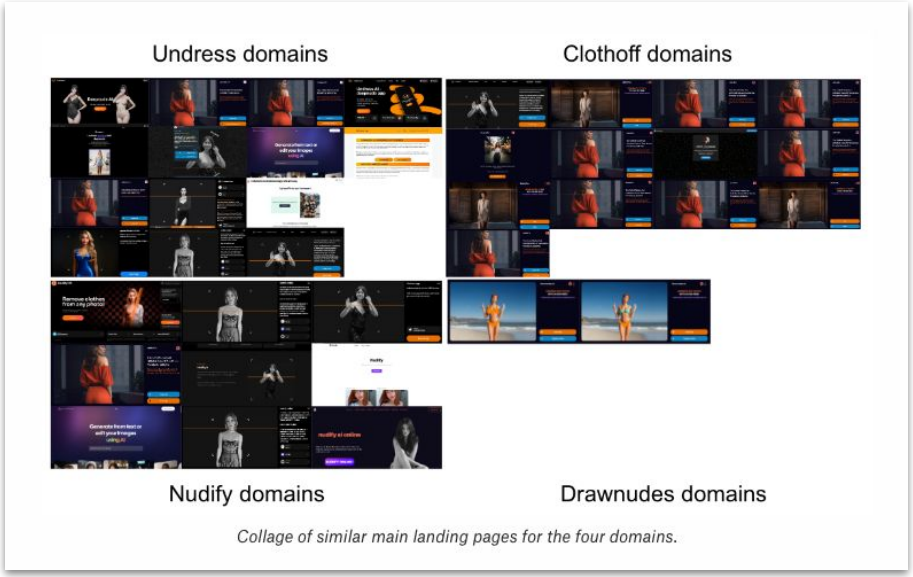
- Accounts on videogame marketplace G2A sold tokens to use on non-consensual pornographic deepfake site
- A network of these sites shared similar TLDs — this is the kind of scenario where dnstwist could be useful
- Sites redirected to each other



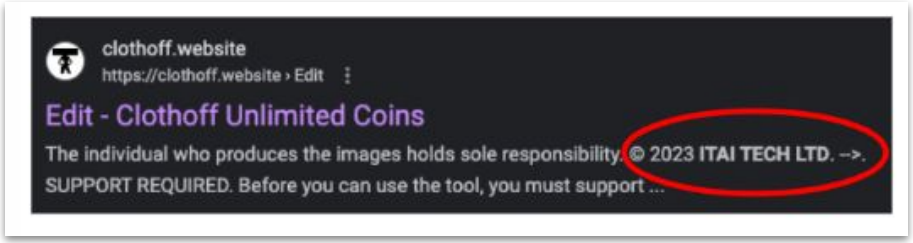
https://www.bellingcat.com/news/2024/02/23/behind-a-secretive-global-network-of-non-consensual-deepfake-pornography

- Looking at all the domains for the sites revealed shared layouts and photographs
- A Google cache* archive of the sites shared a common company listed
- Payment pages for each site all shared the same design and template
- An analysis using RiskIQ revealed a dedicated server IP for multiple deepfake sites

* Google just removed all cached pages from search results 🤷‍♂️

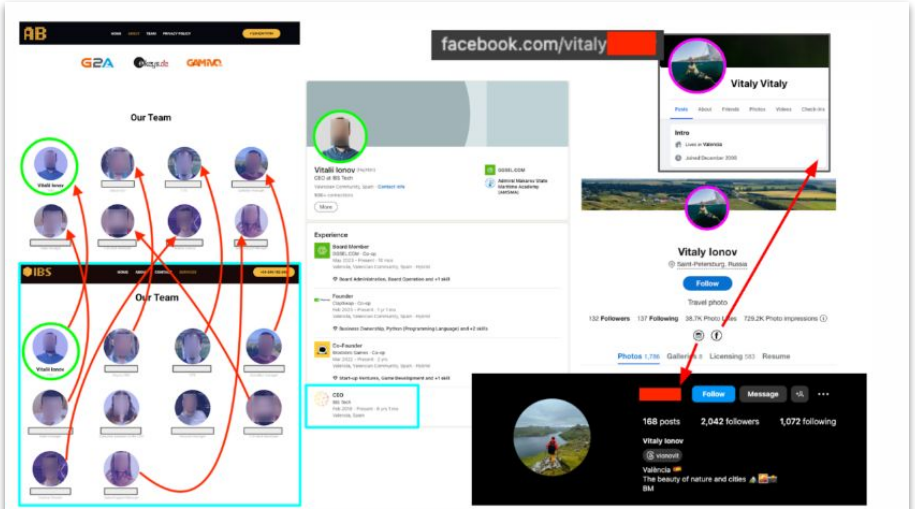


Collage of similar main landing pages for the four domains.



https://www.bellingcat.com/news/2024/02/23/behind-a-secretive-global-network-of-non-consensual-deepfake-pornography

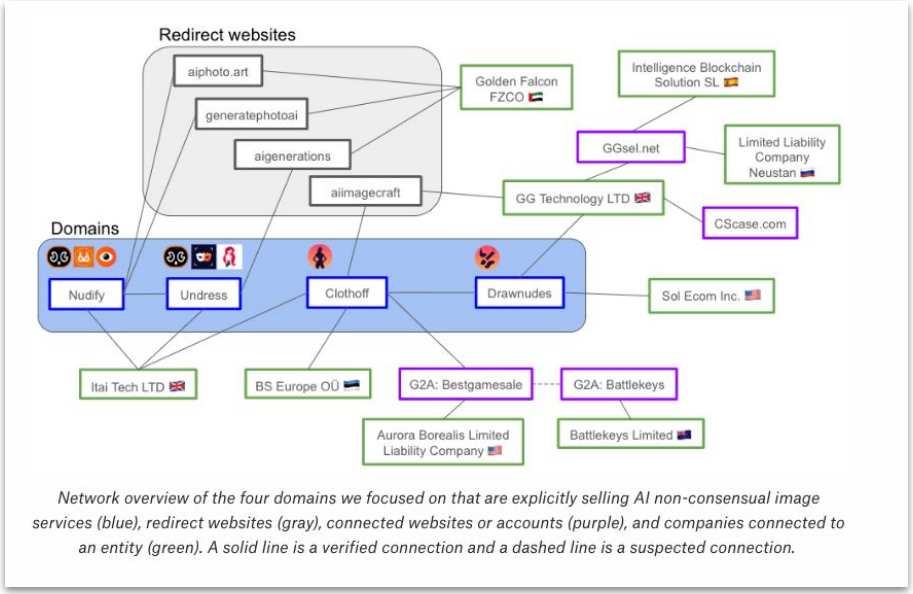
- Examining business registrations found connections between companies linked to the deepfakes sites
- Using “dealers” as middlemen to broker token transactions on mainstream payment platforms, the companies were able to evade detection
- Telegram bots prompted buyers how to jump through hoops to process payments



Screenshots connecting Ionov’s social network accounts and the team pages for Aurora Borealis and IBS Tech.

https://www.bellingcat.com/news/2024/02/23/behind-a-secretive-global-network-of-non-consensual-deepfake-pornography

- Taken together these connections provide strong evidence of a network of entities seeking to obscure its owners, disguise transactions and evade detection
- *"However, these organisations have not perfectly covered their tracks. There are real people who have lent their names, and even their companies, to these platforms."*



Network overview of the four domains we focused on that are explicitly selling AI non-consensual image services (blue), redirect websites (gray), connected websites or accounts (purple), and companies connected to an entity (green). A solid line is a verified connection and a dashed line is a suspected connection.

https://www.bellingcat.com/news/2024/02/23/behind-a-secretive-global-network-of-non-consensual-deepfake-pornography

Resources

Newsletters

Craig Silverman's Digital Investigations

Digital Digging with Henk van Ess

Jack Creps' The OSINT Newsletter

Excellent books

Verification Handbook edited by Craig Silverman

Michael Bazzell's Open Source Intelligence Techniques

https://bit.ly/nicar24-behind-this-website

https://github.com/jonkeegan/behind-this-website

Thank you!

Jon Keegan
Investigative Data Journalist
The Markup
@jonkeegan@mastodon.social

Priyanjana Bengani
Senior Research Fellow
Tow Center for Digital Journalism
@acookiecrumbles@indieweb.social