The Web of Cognitive Warfare

Ву

Jindrich Karasek aka 4n6strider





TLP:GREEN / All slides

Adversary: Syndicated cyber crime gangs and the APT actor

Capabilities:

Data mining, sociomapping, microtargeting, use of doppelganger media pages, algorithm-driven content feeds.



Infra:

Social media platforms, marketing tools, website traffic analysis, cloud-based data storage, and trackers.

Victim: General public, in "enemy" countries

Plan:

The Big picture

OODA: Observe

OODA: Orient

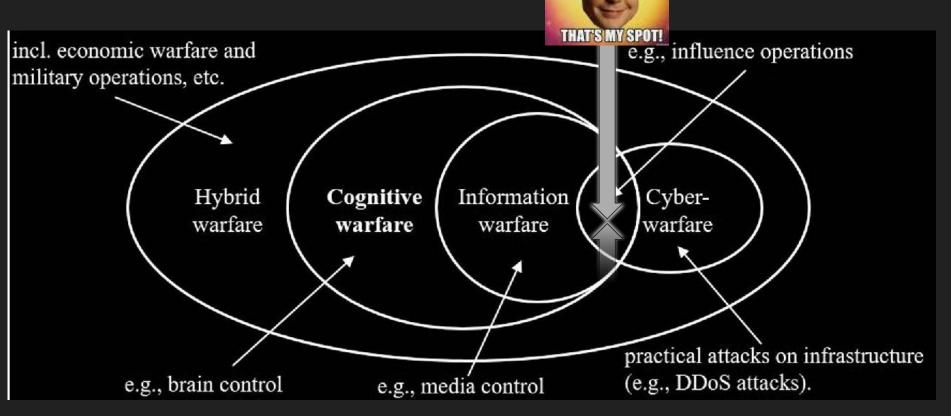
OODA: Decide

OODA: Act

Conclusion



Big picture:



Source: https://academic.oup.com/jogss/article/7/4/ogac016/6647447



ACT

Deploying countermeasures and protective actions Launching public awareness campaigns Utilizing ethical hacking to test and strengthen defenses



OBSERVE

Continuous monitoring of social media and web traffic Identifying indicators of manipulation (e.g., sock puppet accounts, paid ads)

Using OSINT for comprehensive threat landscape awareness



DECIDE

Developing targeted countermeasures Crafting communication strategies to counter misinformation

Planning educational initiatives to raise awareness



ORIENT

Sociomapping to understand network influences

Analyzing data for behavioral patterns and relationships

Identifying cognitive biases and manipulation tactics

Sentiment analysis to gauge public sentiment Named Entity Recognition (NER) to identify key actors

Topic Modeling to uncover hidden themes



OODA 1/4: Observe

Continuous monitoring of social media and web traffic, Identifying indicators of manipulation (e.g., sock puppet accounts, paid ads), Using OSINT for comprehensive threat landscape awareness

Why the Web?

It is everywhere. Good.



UX dark patterns

Manipulative design tactics.

Misinformation loops
Echo Chambers
Information Overload
Fear Appeals
Intermittent Variable Rewards
False Consensus Effect
Hidden Information and Choices

=> Recognition of dark patterns is essential in safeguarding democratic integrity and promotion of informed decision making.



<u>Phenomena</u>

- {Explained in Appendix}
- Narrative
- Overton Window
- Meme
- Fear
- Cognitive biases
- Persuasion
- Dark Patterns in UX
- Social Identity
- Group Thinking
- Decisions by emotions
- Belief change
- Cognitive capacity



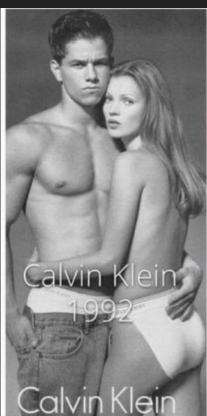
[Caution: This slide is all hoaxes!]

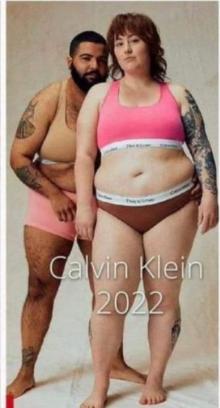
5 Embassy of Russia in Bangladesh

2 d ⋅ 3

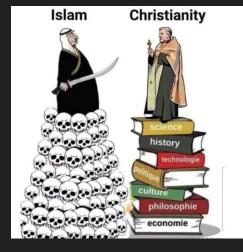
Alexander Dugin (https://t.me/Dugin_Aleksandr): At the heart of the global confrontation that has begun is the spiritual,... Zobrazit víc



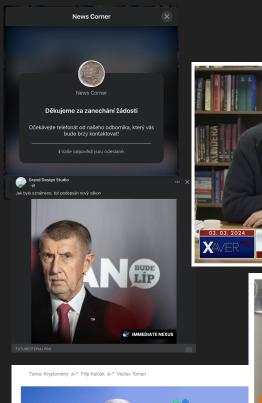








Ongoing PsyOP in Czechia:









Šéf kontrarozvědky Michal Koudelka sdělil o úspěšné operaci na chytání internetových podvodníků. České úřady zahájí rychlé odškodnění všem poškozeným v důsledku činnosti podvodníků

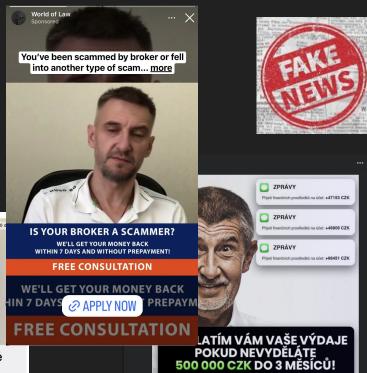




Viděli jste kurz bitcoinu? Jsem si jistý, že nyní je ten správný čas investovat do kryptoměn. A já už jsem to na své finanční platformě Immediate Matrix udělal.



210 Comments 103 Shares



OpenAl

1) Paid Ads promotion:

Social media:
Use of doppelganger media pages





1) Paid Ads promotion:

Social media: Use of doppelganger media pages

Build: Promoted by sock puppet accounts - comments, likes





1) Paid Ads promotion:

Social media: Use of doppelganger media pages

Build:
Promoted by sock puppet accounts - comments, likes

Algorithm starts to feed the user more of related content, conspiracies and "soft" geopolitical content.





1) Paid Ads promotion:

Social media: Use of doppelganger media pages

Build:

Promoted by sock puppet accounts - comments, likes

Algorithm starts to feed the user more of related content, conspiracies and "soft" geopolitical content.

Cognitively influenced audience feels that this is natural cause of action and are willing to recommend to others.





1) Paid Ads promotion:

Social media: Use of doppelganger media pages

Build:

Promoted by sock puppet accounts - comments, likes

Algorithm starts to feed the user more of related content, conspiracies and "soft" geopolitical content.

Cognitively influenced audience feels that this is natural cause of action and are willing to recommend to others.

2) Data mining:

Harvesting of the data using the forms, scripts and usual marketing tools and website traffic analysis





1) Paid Ads promotion:

Social media: Use of doppelganger media pages

Build:

Promoted by sock puppet accounts - comments, likes

Algorithm starts to feed the user more of related content, conspiracies and "soft" geopolitical content.

Cognitively influenced audience feels that this is natural cause of action and are willing to recommend to others.

2) Data mining:

Harvesting of the data using the forms, scripts and usual marketing tools and website traffic analysis



3) Data processing:

Data are sent to third party servers in Cloud, javascripts use often trackers of Yandex, Facebook or Google





1) Paid Ads promotion:

Social media: Use of doppelganger media pages

Build:

Promoted by sock puppet accounts - comments, likes

Algorithm starts to feed the user more of related content, conspiracies and "soft" geopolitical content.

Cognitively influenced audience feels that this is natural cause of action and are willing to recommend to others.

2) Data mining:

Harvesting of the data using the forms, scripts and usual marketing tools and website traffic analysis

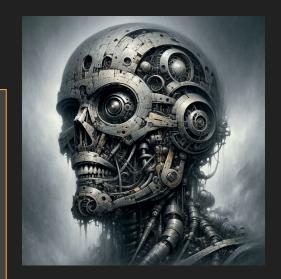


3) Data processing:

Data are sent to third party servers in Cloud, javascripts use often trackers of Yandex, Facebook or Google 4) Output:

The actor has detailed personal info about audience, their relations and has a financial gain if the scam succeeded.









https://www.disarm.foundation/framework https://disarmfoundation.github.io/disarm-navigator/



OODA 2/4: Orient

OSINT, mapping to understand network influences, Analyzing data for behavioral patterns and relationships, Identifying cognitive biases and manipulation tactics, Sentiment analysis to gauge public sentiment, Named Entity Recognition (NER) to identify key actors, Topic Modeling to uncover hidden themes



□ Všechny reklamy ▼



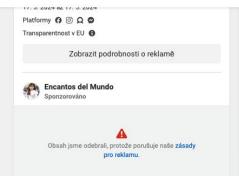


X



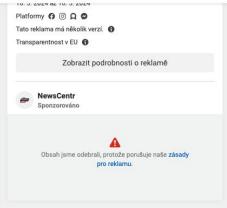
Uložená hledání

Spuštěno: květen 2024









Doručování reklam v EU

Dosah

11 722

Zobrazit podrobnosti o reklamë

ID knihovny: 2438260156385237

Neaktivní

Máte poslední sanci

15. 5. 2024 až 17. 5. 2024

WWW.MATCHOESUCCESSTRNEWS.COM

Platformy () O Q O

2 reklamy používají tento obsah a text

Zobrazit souhrnné podrobnosti

ID knihovny: 974707584354146

Neaktivní

Další inform..

16. 5. 2024 až 16. 5. 2024

Platformy () (Q Q

2 reklamy používají tento obsah a text

Zobrazit souhrnné podrobnosti

Věk 0

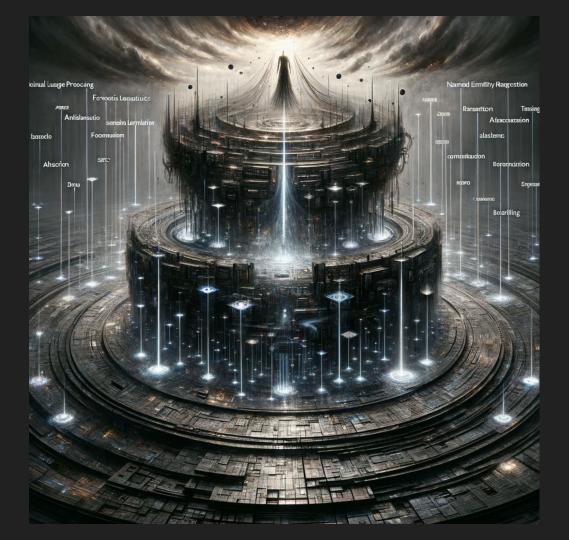
35 - 65 + let

Zobrazit podrobnosti o reklamě

Narrative intelligence: 4 Al/ NLP Tricks:

- NER
- POS
- Topic Modeling
- Sentiment Analysis

Automation saves time and protects the sanity of analysts.



Al comes to save the day. Well, the NLP.

- Sentiment analysis => hate speech detection

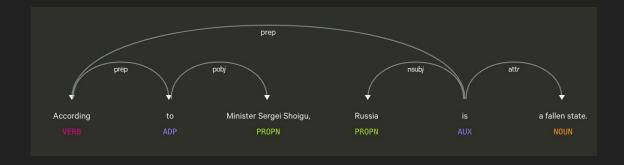
@JohnPowe14 Putin a criminal! @Russia do you support that? Stand with @ukraine! #StandWithUkraine {'polarity': -0.625, 'subjectivity': 0.55, 'assessments': [(['criminal', '!', '!'], -0.625, 'o.55, None)]}

Detection of the named entity => what's going on

Russian NORP President Vladimir Putin PERSON has repeatedly declared that "the Ukrainian NORP operation has failed." According to Russian NORP Defense Minister Sergei Shoigu PERSON , Moscau OPE 's losses from the start of the offensive to the end of October DATE exceeded 90,000 people.

- POS: parts of speech:

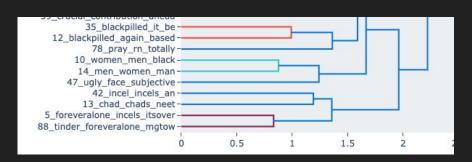
4 tricks you should know about

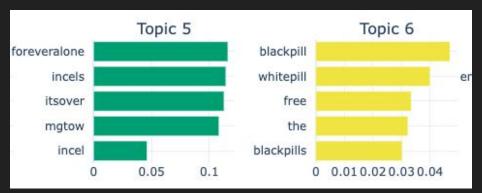


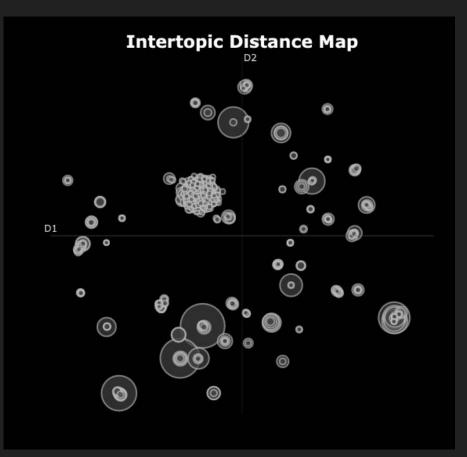
AI TRICK: Topic Modeling

{Topic modeling with Google BERT}

$$\phi("king") - \phi("man") + \phi("woman") = \phi("queen")$$

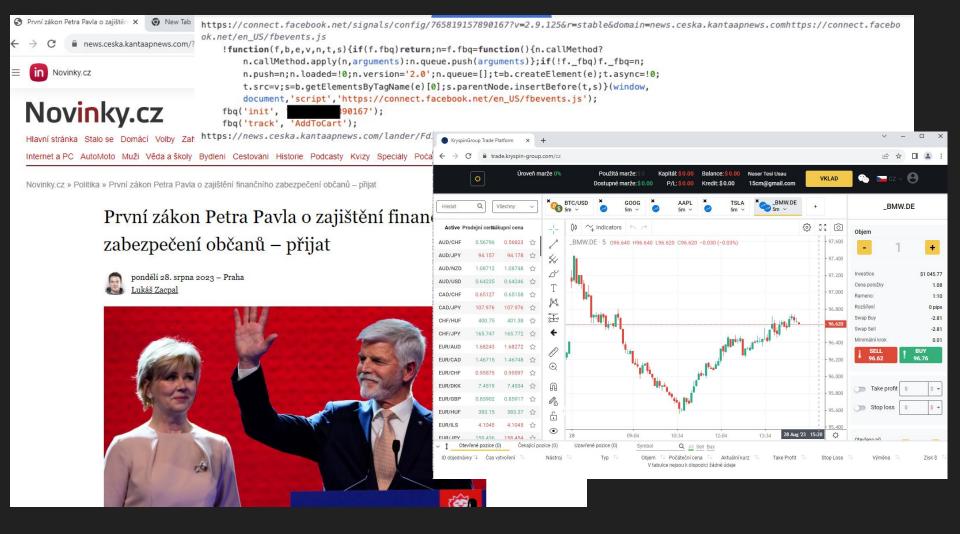






Ethical hacking











first_name is required

last_name is required

email is required

phone must be a valid phone number

+420 + 601 123 456

Poslat zprávu



```
Sources Network Performance
<!DOCTYPE html>
<html lang>
                                             s wafw00f goslakinvest.pro
v <body id="body">
 ▶<header class="page-header"> ··· </header>
  w < main class="page-main">
   v<section class="page-start page-section">
     v < div class="container">
                                                                       Woof!
       ▶<h1 class="page-title
                          text-white text-cent
       ▼ <div class="row align-items-center
                           justify-content-between
        ▼ <div class="col-12 col-md-6 col-lg-5 mt
          v<div class="page-form" id="page-form"</pre>
           w<form action="thankyou.php" method=</pre>
               <input name=" form id" type="hid
             ▶ <div class="page-form group focu
             ▶ <div class="page-form_group focu
                                                                       ~ WAFW00F : v2.2.0 ~
             ▶ <div class="page-form_group focu
                                                  The Web Application Firewall Fingerprinting Toolkit
             ▶ <div class="page-form group"> ....
               <button type="submit" class="page
              <input type="hidden" name="flow_h[*] Checking https://goslakinvest.pro</pre>
               <input type="hidden" name="landing+]</pre>
                                                  The site https://goslakinvest.pro is behind Cloudflare (Cloudflare Inc.) WAF.
               <input type="hidden" name="facebo</pre>
                                                  Number of requests: 2
               <input type="hidden" name="click"[~]</pre>
               <input type="hidden" name="sub1" value="{sub1}">
               <input type="hidden" name="sub2" value="{sub2}">
               <input type="hidden" name="sub3" value="{sub3}">
               <input type="hidden" name="sub4" value="{sub4}">
               <input type="hidden" name="landing_url" value="https://goslakinvest.pro/registration/"> == $0
                                                            [kali⊕kali)-[~]

$ sudo nmap -sS -p- -Pn gclab.ru
        ><div class="col-12 col-md-6 col-lg-7">...</div>
                                                           Starting Nmap 7.94 (https://nmap.org) at 2023-08-22 14:01
                                                           Nmap scan report for gclab.ru (31.10.5.10)
    </section>
   > <section class="features"> - </section>
                                                           Host is up (0.022s latency).
   ><section class="about"> -</section>
                                                           rDNS record for 31.10.5.10: srv17401.hosted-by-eurohoster.or
   </main>
  ▶<footer class="page-footer"> -</footer>
```

<script defer src="js/bundle.741f2dd...js"></script>

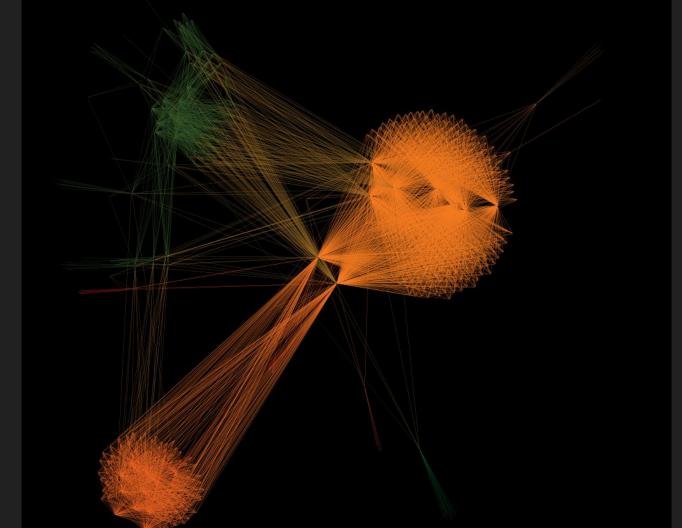
<script src="js/landing url2.js"></script>

</body>

</html>

```
▶<div class="modal" id="privacy" tabindex="-1" aria-labelled Not shown: 65528 filtered tcp ports (no-response)
 <script type="text/javascript" src="https://posthog.gclab.ru PORT</pre>
                                                             STATE SERVICE
 <script defer src="assets/errors-modal/js/errors-modal.js">
                                                 22/tcp
                                                             open
                                                                     ssh
                                                 80/tcp
                                                                    http
                                                             open
                                                                   https
                                                 443/tcp
                                                             open
                                                 1194/tcp closed openypn
                                                 5001/tcp closed commplex-link
                                                 19000/tcp open igrid
                                                 19001/tcp closed unknown
```

Nmap:



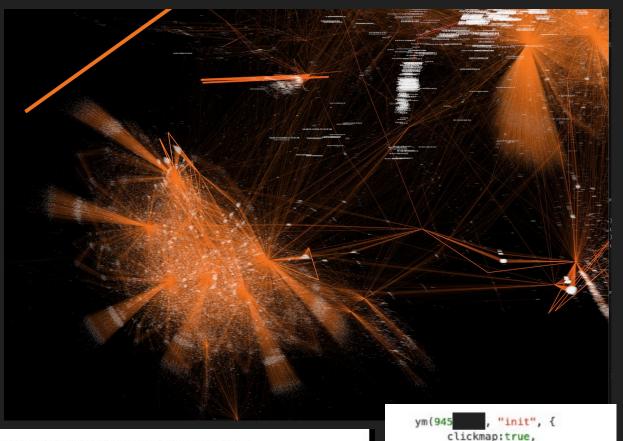
Infra / Scripts: Js, PHP..

```
https://geekstoyspalace.com/?sub
https://ggcascadingfalls.site/
https://gostumble.net/
https://govtjobsstudy.com/
https://haramadzha.site/
https://hawa-taps.net/
https://hawksquadservice
                           et/
https://heylink.me/
https://heylink.me/canta
                            ac/
https://heylink.me/casta
                           at/
https://heylink.me/ceaso
                            ie/
https://heylink.me/detal
                           iz/
https://heylink.me/dhir:
https://heylink.me/emie:
https://heylink.me/fral
https://heylink.me/frrorowa
https://heylink.me/jeazipin
```

https://hibabeaute.net/

https://hollvchristmasgift





trackLinks:true,

https://goslakinvest.pro/assets/dat

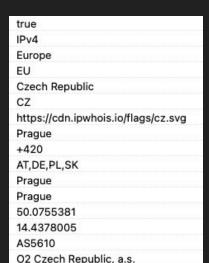
t="" /></div>

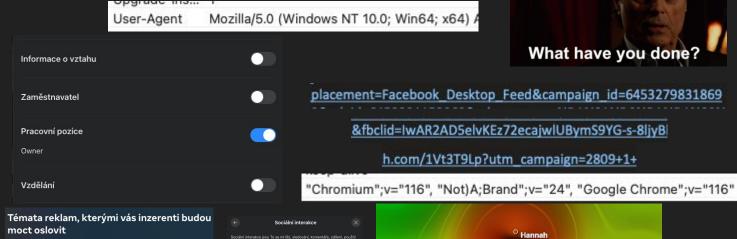
accurateTrackBounce: true

```
2d1
{"ip"
                        "success":true, "type": "IPv4", "continent": "Europe", "continent code": "EU", "country": "Czech
Republic", "country code": "CZ", "country flag": "https:///cdn.ipwhois.io//flags//
cz.svg", "country_capital": "Prague", "country_phone": "+420", "country_neighbours": "AT,DE,PL,SK", "region": "Prague", "city": "Prague", "latit
ude":50.0755381, "longitude":14.4378005, "asn": "AS5610", "org": "02 Czech Republic, a.s.", "isp": "02 Czech Republic,
a.s.", "timezone": "Europe\/
Prague", "timezone_name": "CEST", "timezone_dstOffset":3600, "timezone_gmtOffset":7200, "timezone_gmt": "+02:00", "currency": "Czech Republic
Koruna", "currency_code": "CZK", "currency_symbol": "K\u010d", "currency_rates": 22.326, "currency_plural": "Czech Republic korunas"}
```

All sorts of data the actor is technically able to collect:

Inzerenti vás můžou kontaktovat na základě témat, která se vám podle nás líbí. Tyhle témata můžete zkontrolovat a sprayovat, čeho chcete





Frank

Relationship status, employer, job title, education, likes & dislikes, what triggers you, what do you share and with whom, what advertisement topic you prefer, your frequent contacts, social dynamics.

vidět víc nebo míň

Referer: https://mokwest.com/? fb=888695386202337&utm_campaign=antonis1500log_ads28 D4FdpiXMvK9xnHidkCKIfC0d0lGRTM

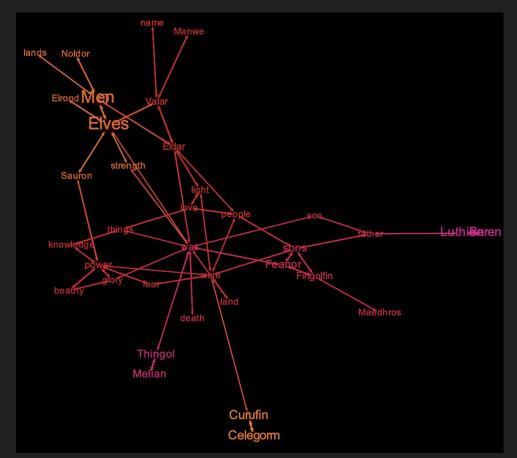
[○] Cyril

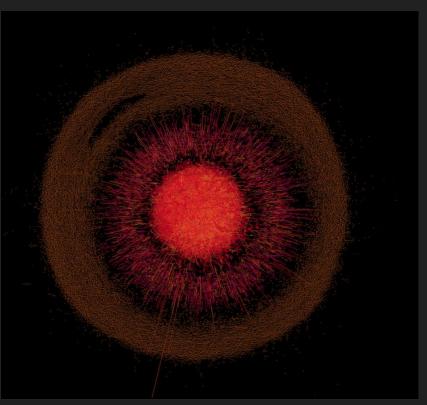
Furthermore:

Email, cellphone, location, bank details, government ID & mailbox

More on Social Graphs / SNA

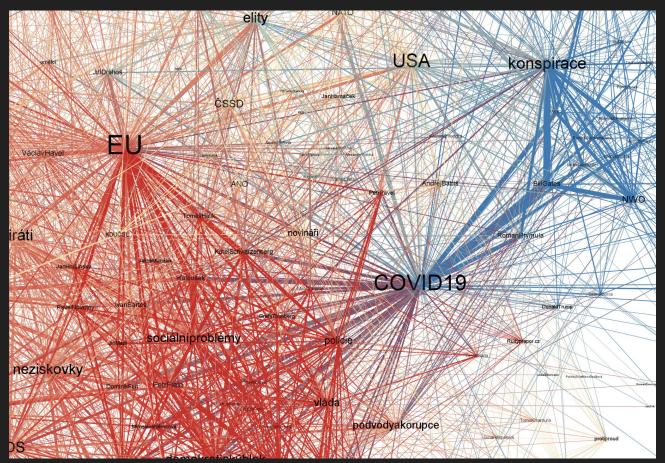
SNA 1/2: Connections of NER by POS to do the "Attribution"

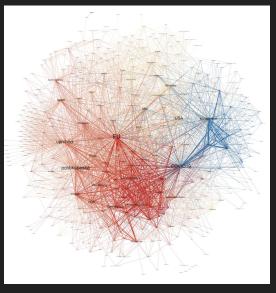




Source: Eolas training datasets based on Silmarillion

SNA 2/2: "Graph" topography:





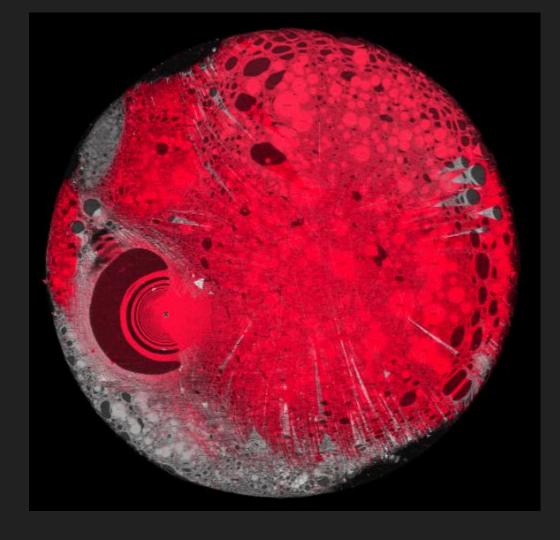
Sources:
Project Eolas Disinformation
domain in CZ in 2020/2021

PsyOps related to Gaza 2024 on X:

Data slice contains Israel / Gaza related contents. {Motivation behind the research is to understand the TTP's of politically motivated threat actors, who often use sophisticated malware, espionage or cognitive warfare. }

- 90% of the content is re-shared by the "influential personas". They may be not always what they seem.
- Red: Calling for cease fire, troops withdrawal, criticising the Israeli actions taken after the terrorists attack. Asking OSN to stop it. Blaming all the civil victims to the Israel. Ooops, no mention about what did started this counter action, why did Israeli entered Gaza.. Not one time! In almost 2 million of mentions.
- White: Mentioning the victims of the terrorists from Gaza.
 Defending the actions of the army of Israel, more fact checking related contents. Both Red and White are heavily re-shared and promoted, meaning both are more than just "authentic user behaviour on a social network".

Data from the period 12-2-2024 to 18-2-2024, 7 days worth data, \sim 1,3 million of new "posts", 873743 nodes and 1916028 edges in the Graph.



OODA 3/4 : Decide

Developing targeted countermeasures, Crafting communication strategies to counter misinformation, Planning educational initiatives to raise awareness

Measure of success of the PsyOps for StratCom:
The Overton Window shift

Concept from the Political science

Describes the range of acceptable ideas in public discourse

Not fixed, shifts over time depending on public opinion

Manipulated introduces and normalizes extreme ideas.



Process map:

TA PoV: Control the narrative + financial gain + social mapping



Paid ads promotion



Data mining



Data transfer + processing



Output

TLP:GREEN

Infrastructure

Meta x Youtube x Web search results

Forms filled in by the visitor, Web or Meta based

Javascript, cloud based servers

Final data receptors, covered by proxy. Maintained over SSH / VPN Detection

Semantic: Stolen or sock puppet account promotes Investments or scandalous relevelation, track the advertiser ID

Data validation javascripts, tracking ID and URL tokens

Netflow + CDN data analysis

Netflow + data collector scripts analysis

Disruption

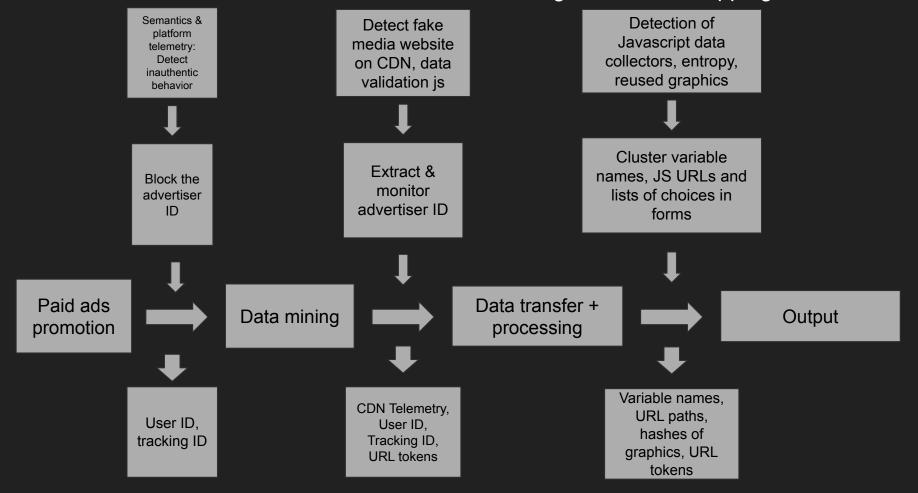
The Platform: Blocking of inauthentic behavior and its monetisation

CDN based detection patterns

Awareness, fraud detection rules

Killchain - TA Goal: Control the narrative + financial gain + social mapping

TLP:GREEN



OODA 4/4: Act

Deploying countermeasures and protective actions, launching public awareness campaigns, Utilizing ethical hacking to test and strengthen defenses

Cognitive Warfare Defense and Security

Strategic advisory, Awareness, Active countermeasures, StratCom support Named Members:

Jindrich Karasek - Cyber and CTI domains, also PhD at FHS UK

Iva Polackova Solcova - Institute of Psychology, Czech Academy of Sciences.

Josef Holy - Faculty of Humanities, Charles University in Prague

Eva Klusova - Faculty of Humanities, Charles University in Prague

Call for action:

To join the hunt within your own domain, country, company .. to protect the data, brand, elections or democracy.

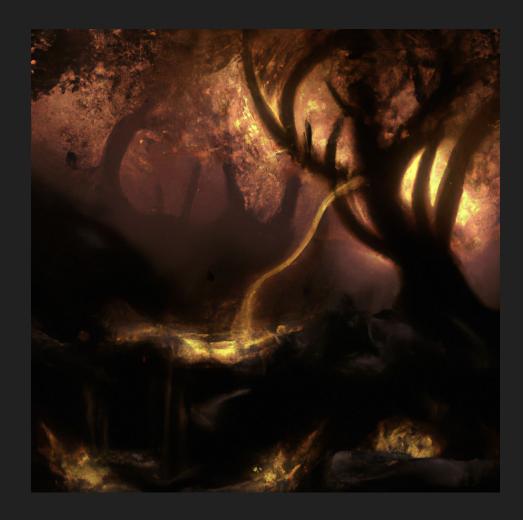
Also, seeking the opportunities for academic publications for my PhD.

Conclusion:

- Cognitive warfare is the most serious threat to our society

 (Because it blocks us to do reasonable decisions..)
- The Mind has no firewall.
- To survive learn and know yourself.

(The only way how to remain valid in fast paced civilisation. Keep evolving and don't get controlled by your fear. Small everyday good deeds everyone capable of doing)



Appendix: Further study links:

How to detect transparent proxy: https://catching-transparent-phish.github.io/catching-transparent-phish.githu

Google fog of war:

https://blog.google/threat-analysis-group/fog-of-war-how-the-ukraine-conflict-transformed-the-cyber-threat-landscape/

DTAC:

https://blogs.microsoft.com/wp-content/uploads/prod/sites/5/2023/02/DTAC-Attribution-Framework.pdf

"Please forward before they delete it!" Assessing conspiracy theories and disinformation in chain e-mails about Covid-19 in the Czech Republic [version 1; peer review: awaiting peer review] https://routledgeopenresearch.org/articles/2-25/v1

ENISA:

https://www.enisa.europa.eu/publications/foreign-information-manipulation-interference-fimi-and-cybersecurity-threat-landscape/@@download/fullReport

NATO Posture:

https://www.researchgate.net/publication/369305190_Mitigating_and_Responding_to_Cognitive_Warfare

DISARM Kill chain model:

https://fighting-fake-news.eu/articles/kill-chain-model-disinformation

Poc Python code for the analysis: (There is also private repo, accessible only upon request) https://github.com/Eolas-bith/UNICRI-Decode-CW

Another activity group of the actor: https://dfrlab.org/2024/03/12/doppelganger-operation-targets-ukraine/

My research notes:

https://www.linkedin.com/pulse/enhancing-defenses-against-cognitive-warfare-through-cyber-karasek-co5sc/

https://www.bellingcat.com/category/resources/

https://www.bellingcat.com/resources/2024/04/25/oshit-seven-deadly-sins-of-bad-open-source-research/

Thank You!

Jindrich Karasek aka 4n6strider

Senior Cyber Threat Researcher



TLP:GREEN / All slides

