ACIG
APPLIED CYBERSECURITY & INTERNET GOVERNANCE

NASK

Cyber Influence Defence: Applying the DISARM Framework to a Cognitive Hacking Case from the Romanian Digital Space

Alina Bârgăoanu | Faculty of Communication and Public Relations, National University of Political Studies and Public Administration, Bucharest, Romania | ORCID: 0000-0003-3912-8442

Mihaela Pană | National University of Political Studies and Public Administration, Bucharest, Romania

Abstract

One of the main lessons learned in the context of Russia's full-scale invasion of Ukraine starting in February 2022 is that foreign information manipulation and interference (FIMI) operations are closely coupled with cyber threats. Regardless of whether cyberattacks are followed by an information manipulation component and vice versa, the merger of the two can be an early indicator of the potential for a conflict to escalate from the cyber area to the ground. Our article is premised on the idea that today's highly technologised information ecosystem is a fertile ground for cyberattacks and information manipulation in the context of FIMI; more specifically, it enables cognitive hacking, meaning hacking the human mind and human cognition altogether through technological disruption and cyber pressure. Starting from this premise, the aim of the article is to highlight the technological determinants of cognitive hacking and identify silent or emerging threats that bypass technological sensors and seek to disrupt and manipulate the information environment. The empirical part is based

Received: 26.03.2024

Accepted: 02.06.2024

Published: 10.07.2024

Cite this article as:
A. Bârgăoanu, M. Pană "Cyber influence defence: Applying the DISARM framework to a cognitive hacking case from the Romanian digital space," ACIG, vol. 3, no. 1, 2024, DOI: 10.60097/ACIG/190196

Corresponding author:
Alina Bârgăoanu, Faculty of Communication and Public Relations, National University of Political Studies and Public Administration, Bucharest, Romania; E-Mail: alina.bargaoanu@comunicare.ro
0000-0003-3912-8442

on observation as a descriptive method, which is used to analyse a case of cognitive hacking carried out via a YouTube malvertising campaign targeting Romanian users. This case study is analysed qualitatively by matching the DISinformation Analysis & Risk Management (DISARM) framework with evidence collected through Open-Source Intelligence (OSINT) tools, following an innovative analysis structured according to the purposes, actions, results and techniques (PART) model. The extensive analysis of the identified case shows that applying the DISARM framework to cyber-enabled operations can be useful for anticipating and responding to FIMI threats, even when such operations do not appear to have a specific, immediately identifiable purpose.

Keywords

cognitive hacking, FIMI, cyberattacks, cyfluence, deepfake, OSINT analysis, DISARM framework, malvertising

1. Introduction

After the COVID-19 pandemic and the accompanying infodemic, humanity reached a flashpoint with two simultaneous geopolitical conflicts that present the potential to disrupt the current world order. Analyses of the events of the last 4 years converge on the thesis that the cognitive dimension has become a new frontier of offensive and defensive military actions. Russia's full-scale invasion of its neighbouring country, Ukraine, coupled with the conflict between Hamas and Israel following the 7 October 2023 terror attacks, led to the new hybrid threat architecture, at the heart of which lies the battle for peoples' minds, enabled by our dependence on technological structures. In this turbulent context, the threat of cyber influence could be disguised as a regular cyber-crime that bypassed technology filters silently and crosses all the adversary lines.

Given the immaterial environment of the human mind, where the effects of hostile actions can only be inferred from people's perceptions, decisions and behaviours, how can cyber interference be proved? Does technology provide the same conditions to track attackers through digital fingerprints and build a behavioural profile to determine the threats against which to protect oneself? The answers to these questions form the basis of this research, which lies at the intersection of information security and communication studies.

To exemplify the theory of cognitive warfare conducted by combining information operations with cyberattacks to enhance psychological effects, the objective of this paper is to describe and analyse a cognitive hacking case using multiple tools and methods, with the aim of consolidating the practice of hybrid threat-integrated anticipation and response. Specifically, by observing two inauthentic video ads on YouTube targeting Romanian users, this paper analyses how deepfake videos, fabricated content and compromised websites are blended together to deliberately spread false information and malware. This case study provides insight into the hybrid approach needed to effectively manage a hybrid threat, such as cognitive hacking, using open-source tools and an innovative strategic analysis framework.

The case study findings lead to the analysis of cognitive hacking by tracking disinformation and malvertising – a method used to describe misleading ads that contain malicious code or redirect users to malicious websites [1, 2]. This case study reveals how to use Open-Source Intelligence (OSINT) for evidence-gathering in the attribution of hostile actions and how to apply the DISinformation Analysis & Risk Management (DISARM) framework to cyber-enabled influence campaigns for anticipating foreign information manipulation and interference (FIMI) operations, even when such operations do not appear to have a specific, immediate identifiable purpose.

1.1 Cognitive Hacking in the Context of the Russia–Ukraine Cyber War

A good understanding of cognitive hacking is related to the large picture of Russian cyber operations aimed at extensively disabling Ukraine's critical national infrastructure [3], telecommunications, banking, transport, water supply and energy supplies [4] during the past 10 years. This concept emerged at the disruptive cyberattacks of the first major crisis in Eastern Europe, the pro-European protests in Ukraine that took place in 2013 under the name EuroMaidan [5], and grew intensively before and after the armed conflict triggered by Russia in Ukraine [6–8], shifting to the human cognitive dimension as a new type of critical national infrastructure [9]. Weaponising the online manipulation capabilities of new technologies [10] and exploiting human addiction to social media, the weak control mechanism of the distribution of online content and undetected technical vulnerabilities create the premises for cognitive warfare [11–13].

1.2 Cognitive Warfare: From Cyber-Enabled Influence to Cyfluence and Cognitive Hacking

From a technological perspective, both humans and information systems can be viewed as the endpoints of information exchanges [14]. According to Cybenko's early research, if influence operations are deliberate activities targeting the cognitive dimension with the aim of changing the attitude or behaviour of the target audience, as Hollis concluded [15], cognitive hacking refers to a computer or information system attack that relies on changing human users' perceptions and corresponding behaviours to be successful [16]. In NATO's approach, cognitive warfare integrates cyber, information, psychological and social engineering capabilities. These activities, carried out in conjunction with other instruments of power, can affect attitudes and behaviour by influencing, protecting or disrupting individual and group cognition to gain advantage over an adversary [17]. New and emerging technologies, such as artificial intelligence (AI) and deepfake, combined with disinformation, microtargeting and algorithmic echo chambers reveal the future of hybrid threats [18].

Seen as a 'strategy that focuses on altering how a target population thinks and through that how it acts' by Backes and Swab [19] and 'the weaponization of public opinion, by an external entity, for the purpose of influencing public and governmental policy and destabilizing public institutions' in Bernal et al.'s findings [20], cognitive warfare is determined by at least two essential components: *cognitive domain operations* (CDOs), which use emerging technologies to advance battles into 'the realm of the human mind' [21], and coordinated chaos [22], which synchronises cyberattacks and disinformation to manufacture crises and disrupt public responses as a 'never-ending battle for minds' [23].

In line with the latest research findings, the approach of treating the cognitive dimension as an offensive and defensive manoeuver space has emerged from the US military [9]. The analysis of Russia's actions over the past 10 years, culminating with the outbreak of a full-scale military invasion in February 2022, reveals the hybrid nature of offensive and defensive actions and the integration of technology in attempts to destroy or weaken the adversary from a cognitive point of view [24, 25].

While analysing the fusion between hostile influence campaigns, cybersecurity and AI, Yonat points out that 'the attackers are light years ahead of us and moving faster than us'. He explains 'cyfluence' as a concept used to define the embedding of cyberattacks in

influence campaigns [26]. He also highlights the cataclysmic effect of using AI in influence operations, 'not just damaging companies or individuals or just harming countries; it is literally tearing apart societies, bringing down democracies, taking humanity one enormous step toward another dark age' [26].

The contemporary information ecosystem has created 'the worst cognitive warfare conditions since WWII' [27], affecting a nation's cognitive infrastructure, which Gourley described for the first time as 'the mental capacities of the citizens and the decision-making ability of people, organisations, and our government' [28]. Regarding responses to this new type of threat, the Swedish approach appears to be the most advanced model. Established in 2021, the Swedish Psychological Defense Agency, organised as a government agency under the Ministry of Defense, is in charge of identifying, analysing and countering foreign malign information-influenced activities [29].

1.3 Convergence between Disinformation, Influence and Cyberattacks

The concept of cyber pressure can be related to the increasing number and sophistication of cyberattacks [30], hybridisation of attackers' motivations and techniques, increased risk of an unknown vulnerability being exploited without any possibility of knowing it, lack of adequate cyber threat anticipation as a result of poor technological knowledge, and poor resource allocation under time pressure, technological illiteracy among users, poor communication skills of technical specialists, the speed of technological transformation, and an unpredictable and unstable geopolitical environment. Given this pressure, cyberattacks have become part of the ecosystem of disinformation operations [31, p. 9], which is why the cyber risk associated with this threat is considered at all levels, from business [32] to national security [33].

The hybridisation of attacks by combining cyber and information warfare to create social harm has a new pattern: cyberattacks are used as a tool for information attacks, and information attacks are used to amplify the alleged success rate of cyberattacks. Both seek to strain people's trust in public action and public entities, create a general sense of insecurity, and erode the capacity to act and react under crisis situations. '[Distributed denial-of-service] DDoS attacks and defacement erode people's trust in their institutions and their ability to protect their own population' [34].

Covert cyber operations are carried out through techniques and tactics, such as social engineering, phishing campaigns, the penetration and capture of computer systems, and the development and control of troll and bot farms. Hacking computer systems to extract documents, publishing illegally accessed documents in truncated or altered versions, capturing legitimate email or social media accounts to disseminate false information, and penetrating content management systems of official websites to spread influence narratives are part of the arsenal of techniques used in cyber influence operations [35, pp. 120–124].

Misinformation and disinformation are recognised among the security threats included in the official analysis of European Union (EU) [36], in direct association with the notion of attempts to influence human behaviour. Furthermore, the European Cyber Security Agency (ENISA) report states that these two threats have become the core of cybercrime activities, which have led to the emergence of the Disinformation-as-a-Service (DaaS) business model. The EU Cybersecurity Strategy [37] also states that hybrid threats combine disinformation campaigns with cyberattacks on infrastructure, economic processes and democratic institutions, with the potential to cause material damage, facilitate illegal access to personal data, facilitate the theft of industrial or state secrets, sow distrust, and weaken social cohesion.

If the main objectives of hybrid warfare are to take control of society, influence people's cognition and disrupt decision-making processes, as well as to gain access to a country's strategic, communication and critical infrastructures by effectively combining soft and hard power [38], then the ability to weaponise new technologies attracts the attention of entities interested in global domination or at least disruption of cyberspace. Researchers have identified the emergence of online influence operations since 2004. As states have shown interest in online influence using microtargeting [39, p. 47], the phenomena of fake news, misinformation and disinformation have become serious challenges to modern society [40]. Consequently, the covert use of social media by promoting propaganda, advocating controversial and toxic narratives, playing both sides of highly divisive issues, and spreading misinformation have become common tools [41].

Analysing how different state actors deployed cyber tools and tactics for hybrid warfare during a major crisis over the past decade, Duggan [42, p. 47] described the 'synchronized choreography' between disinformation and cyberattacks, which can help people gain time and space for conventional military forces. The ability to

penetrate the computer systems of individuals, organisations and institutions significantly increases the potential for effective disinformation and propaganda delivered through both traditional and unconventional means. Thus, cyber actions can increase the potential of influence operations and enrich the information content available to information warfare operators. Cyberscale operations also have socio-psychological effects on citizens and security institutions by distracting attention from the broader manifestations of information warfare [43, 44, p. 12].

The toolkit of hostile actions enabled by the highly technologised cyber environment has grown in variety and sophistication: false information, hyper-partisan content, disinformation, impersonation, false identities, trolls or bot farms, deepfakes, cheapfakes, hacking, hijacking, disconnecting or destroying mobile devices, stealing sensitive information, and leaking personal data. All these hostile actions are encompassed under umbrella concepts, such as cyber-enabled foreign interference [45] or cyber-enabled information warfare and influence operations [46], associated with tools of hybrid interference [47] or forms of hybrid warfare [48].

Zurko, a cybersecurity researcher at MIT Lincoln Laboratory, argues that

> in cybersecurity, attackers use people as a means to undermine a technical system. Disinformation campaigns are designed to impact human decision-making; they are the ultimate use of cyber technology to undermine people. (...) Both use cyber technology and people to achieve a goal. Only the goal is different. Just like cyberattacks, influence operations often follow a multistep path, called a kill chain, to exploit predictable weaknesses [49].

For this reason, Lincoln Lab's efforts are focused on '*source tending*' as well as strengthening the early stages in the kill chain to find new countermeasures for disinformation campaigns.

The ENISA and the European External Action Service (EEAS) have underlined the link between disinformation and cyberattacks and focused on the concept of FIMI. This concept is included in the cybersecurity threat landscape [50] and is used to describe a largely non-illegal pattern of behaviour that threatens or has the potential to negatively impact political values, procedures and processes. Such activity is manipulative in nature and carried out in an intentional and coordinated manner. Additionally, the misinfosec conceptualised by

Walker [51] brings forth the idea of using an information warfare kill chain to understand cyber-enabled influence operations. For this reason, the DISARM framework [52] organises ways of describing and analysing disinformation, covering intent to deceive, intent to harm, and coordinated inauthentic behaviour.

Developed based on cybersecurity best practices, the DISARM framework is designed to gain a common understanding of digital disinformation. The project was designed to codify and share intelligence on disinformation and influence operations through a knowledge base of techniques and countermeasures and presented as a standard that the EU and the United States are now using to analyse and share information in countering FIMI threats [53].

The DISARM phases refer to the highest-level grouping of tactics and their associated techniques, corresponding to a specific time interval in the execution of an influence campaign [54]. If a tactic reveals the adversary's goal for each stage, the techniques lead the way in which the goal is achieved. The kill chain represents the minimum number of steps required for a successful attack. A broken link results in a failed attack, which is beyond the scope of tagging research. Following the DISARM approach, this paper tests the frameworks to identify a case of cognitive hacking from a cyber-enabled influence campaign [55].

2. Methodology

Building upon this conceptual framework, we delve into a detailed examination of a malvertising campaign to investigate how advertising platforms can be utilised for cognitive hacking. To accomplish this goal, our case analysis demonstrate the utility of open-source information in identifying the tactics, techniques, and procedures (TTPs) employed in a cognitive hacking campaign and how these can be matched within the DISARM framework to counter FIMI operations.

The empirical part is based on a case study as a descriptive method that allows for a detailed understanding of a particular case. The analysis is focused on YouTube advertising campaigns targeting Romanian users by showing how deepfake videos, fake news, and compromised websites are blended to deliberately spread false information and malware.

The cognitive hacking case was first spotted on YouTube in May 2023, running in two YouTube video ads about some benefits for

vulnerable social groups without specifying who offers them and under what conditions they can be obtained. The obvious pattern of misleading based on inauthentic content suggests that a large-scale malicious campaign that needs to be captured and investigated before any efforts are made to remove it from the online space. For further reference to this case, we call it YouTube_benefits_Ro.

For this case analysis, the research strategy involves five steps: case identification, message analysis, digital analysis, and OSINT analysis – to track digital fingerprints and collect evidence of misleading actions following an innovative strategic analysis structure by proposing the purpose, actions, results, and techniques (PART) model. Tagging the technological determinants of cognitive hacking into the DISARM framework contribute to a better understanding of the behavioural profile of this case. The research stages were as follows (Fig. 1):

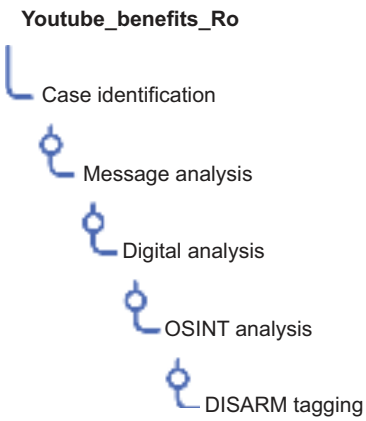


Figure 1. The research stages.

1. Case identification – capture the facts when they happen
2. Message analysis – follow the model of the structured analytic framework based on Lasswell's communication formula
3. Digital analysis – gathering elements related to identified facts
4. OSINT analysis for tracking FIMI fingerprints – collect evidence of misleading actions following the PART model strategy
5. Tagging TTPs into the DISARM framework – version 1.3

3. Research Results

3.1. Case Identification

Video ads targeting vulnerable social groups in Romania were observed when accessing YouTube by nonpaid users in

May 2023. Using a fabricated news flash, an unknown TV presenter announced new benefits of between €5,000 and €10,000 for vulnerable people without mentioning any real recognisable Romanian entity. The targeted audience was mentioned in the second sentence of the message: 'The retired people, pregnant women, low-income people, people with disabilities and many other categories', usually associated with vulnerable social groups with poor cyber hygiene or media literacy to be aware of cyber threats or influence activities. Another misleading clue was the domain of the website mentioned in the video ad, which was redirected to another website.

The high level of uncertainty, an unidentifiable entity, an inauthentic figure, irrelevant visual elements for the audience, and redirection to some subdomains of foreign sites were the triggers to capture this piece of deliberate mislead as it was unfolding and to start the analysis. After refreshing the same page, another video ad stood out, with another presenter and another website related to the ad, but with the same message and the same visual elements.

First capture of the video ad on YouTube

Second capture of the video ad on YouTube

Photo of the video ads on YouTube (ANNEX 1).

After capturing the website and the video, the case was reported to the Romanian National Cyber Security Directorate as an instance of misleading content related to compromised websites. As a result, the sites mentioned above were blocked from being accessed from Romania immediately after that notice.

3.2 Message Analysis

The message analysis follows the model of the structured analytic framework based on Lasswell's communication formula for providing an understanding of the influencing attempts [56, p. 5]. The message was composed of seven short sentences with many unspecified details and unidentified entities expressing supportive behaviour in a polite manner. The only precise elements were the audience – 'retired people, pregnant women, low-income people, people with disabilities and many other categories', the value of benefits – 'planned to be between 5,000 and 10,000 Euro per person', and the call to access the news website (Table 1).

Table 1. The message structure of the video ad promoted on YouTube.

1. Starting this Monday, (unintelligible) introduces benefits for several categories of citizens.
2. Retired people, pregnant women, people on low incomes, people with disabilities and many others will receive benefits.
3. The benefits are planned to be between 5,000 and 10,000 Euro per person.
4. More information can be found on our news website.
5. The method to get the benefits is simple and anyone can do it.
6. You can also read more interesting news.
7. Have a nice day!

3.3. Digital Analysis

The digital analysis is based on public information included in websites promoted in YouTube ads, hhx.theteachingmentors.com and gute.mycalculat.com, to determine as much information as possible about the entity behind the ad campaign and the promoted sites. To perform digital analysis, four actions (A) were carried out.

The first action (**A1**) involved searching the YouTube ad transparency database by website name using the https://adstransparency.google.com/?region=RO tool. The search indicated that the Google Ads Transparency Center has no public evidence of this video

advertising campaign, even though it had been active for at least 3 weeks.

The second action (**A2**) involved checking the websites mentioned in the video ad: hhx.theteachingmentors.com and gute.mycalculat.com. The findings indicate that the websites have the same sitemap: the homepage, one article, and the policy page. There are no active links from homepage to article page, only sensitive images (namely, visualisations of older people in poverty, mentioning safety retirement income, and social security reform) and click bait titles redirected to homepage. All the websites share the same web design, sitemap, and policy page, which is an indication of mass-created websites and a clue that helps to detect and block scam websites used by masquerading attacks. The sites under analysis appeared to be compromised by attackers, as indicated by the fact that they displayed error pages or bad connections during the analysis.

During the third action (**A3**), we checked the content of the websites hhx.theteachingmentors.com and gute.mycalculat.com. There was no information about the data, authors, contacts, or copyrights that could be linked to a real identity.

Finally, the fourth action (**A4**) involved checking the policy page found at hhx.theteachingmentors.com and gute.mycalculat.com. The website privacy policy mentions the Russian Federal Law on Personal Data No. 152 FZ, suggesting that the section is copied from a Russian website. Additionally, this page mentions the name Mihailov Ivan Sergheevici as a data operator (screenshots of the digital analysis are displayed in ANNEX 2). This final evidence helped to discover other websites used in this cognitive hacking campaign during the OSINT analysis stage.

3.4. OSINT Analysis

To perform OSINT analysis for tracking fingerprints and gathering evidence of misleading and harmful actions, we structured the research steps according to the PART model strategy that can be replicated in future OSINT analyses.

The PART model organises the actions (**A**) around the main purposes (**P**) using different OSINT tools for each purpose. The results (**R**) reveal evidence of misleading and harmful actions that can be associated with tactics, techniques, and procedures – TTPs (**T**) – or indicate new directions for analysis purposes. Furthermore, the

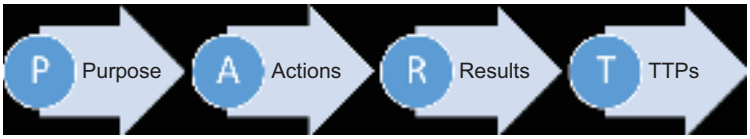


Figure 2. Illustration of the PART model for OSINT analysis of cognitive hacking, as proposed by the authors during the research.

identified TTPs can be correlated with indicators from other databases, such as the DISARM framework, which is described in the next section of the research. Screenshots of the OSINT results are shown in ANNEX 3.

The first purpose (**P1**) was to check for additional information about the websites to which the campaign was leading by performing the domain name search in search engines (**A1**). The Google search results led to one more video ad recorded by a Reddit user (**R1**) that included the essential element – the original YouTube channel that managed the video ad campaign – which has an anonymous and generic name (**T1**): '*România astăzi*' (Romania Today) @romaniaastazi-zl2pj and the evidence of using fake news planted on a newsfeed website weeklynewsfeed.com (**R2**). The fake article planted on weeklynewsfeed.com mixed false information with excerpts copied from an authentic news website (**T2**), including real names of several public officials talking about the Student Invest and Family Start social funding programs and loan facilities of up to €10,000 with interest paid in full by the state. The second video captured by the other Reddit user leads to another website domain name: quoxc.moneyflowgroup.com (**R3**). The analysis revealed that hiding fabricated news in an anonymous newsfeed service is an information laundering technique.

The second purpose (**P2**) was to check the authenticity of the visual content. The video footage shows an official building leading up to an authority representation. Google image identification (**A2**) matches this image with the Ak Orda Presidential Palace in Kazakhstan (**R4**). The correlation of the presenter's physiognomy with the lack of coherence between facial gestures and speech in Romania indicated the use of an AI-generated voice-over for a stock video (**T3**). For this reason, we checked the video with deepfakedetector.ai (**A3**). The result shows a very high probability of deepfake content (**T4**): 71.19% (**R5**).

To complete the third purpose, we checked for any YouTube-related information (**P3**) by performing a thorough search on YouTube.

com (**A4**). The findings indicated that the video ad named '*Pentru cetățenii români*' ('For the Romanian citizens') was posted on 6 May 2023 by the *România astăzi* (România Today) channel and reached over 3.65 million viewers and received 2.1K like reactions and 23 comments (figures from 22 May 2023) (**R6**). The YouTube channel ID @Romaniaastazi-zl2pj has 4.22K subscribers (on 22 May 2023, the day of capture) who joined YouTube on 4 May 2023 (**R7**). By searching for the original video on YouTube, we found that the ad was erased from the initial channel playlist, but it was running as a loop video into a low-profile user playlist. This finding has two meanings: it is a technique used for hiding a video in a shuffle playlist (**T5**) or it is a simple fingerprint generated accidentally by an inexperienced YouTube user. The video ad named 'For the Romanian citizens' was identified in the playlist of user @peisaj131 (URL: https://www.youtube.com/@peisaj131) (**R8**). In this playlist, the video keeps the initial owner names that appear to be the channel named 'Romania Today' – @romaniaastazi-zl2pj (URL: https://www.youtube.com/@romaniaastazi-zl2pj). The profile picture was the evidence of using this channel to manage the video ad campaign (**R9**). At the time of writing this paper, the @romaniaastazi-zl2pj channel was changed to @EvelynTraders – Evelyn Morgan, located in the United States, which shares many videos about FOREX trading to make money easier (**R10**). Meanwhile, the channel has reached 6.86K subscribers (URL: https://www.youtube.com/channel/UCWXYuujcE4lw_JCaLxHeU9Q).

The next purpose (**P4**) refers to finding additional information about the content of the policy page by performing a Google search. With the name 'data controller', Mihailov Ivan Sergheevici (**A5**), two other sites with the same model privacy policy page in Romania, were identified: Kishoregoldsmith.com and pineridgedevelopers.com **(R11)**. The technique identified shows the use of a fake privacy policy (**T6**), an automated translation with some Russian legal references included, without any relevance of data protection of Romanian audience/users.

To explore the website history (**P5**), we checked the Archive.org database for all the websites related to the campaign: hhx.theteachingmentors.com, gute.mycalculat.com, quoxc.moneyflowgroup.com, kishoregoldsmith.com and pineridgedevelopers.com (**A6**). All the sites appear to be compromised or captured by attackers (**R12**). They displayed error pages or bad connections during the analysis, and some of them appeared to have no records, while some of the captured pages were deleted from the tracking records. Thus, the technique identified is that of erasing public records of digital

fingerprints as part of information laundering (**T7**). The next step was checking for Google indexing websites (**A7**) to determine the history of the website on the Internet. During this stage, we discovered that all websites appeared to have a history of at least 2 years and were not created only for this campaign (**R13**). This information led to the technique of using comprised or captured websites (**T8**).

The next purpose is to check for digital identity (**P6**) by searching for domain and subdomain names in multiple databases: who.is, whois.com, subdomains.whoisxmlapi.com, and criminalip.io (**A7**). According to our findings, the domain names of all identified websites had the same name servers in the same class C subnet (the first three numbers of their IPs were identical), meaning that the websites were hosted and managed from the same place (**R14**). Using the WhoisXML API subdomain search tools, it appeared that the subdomains used in this campaign were created between 19 May and 23 May 2023 (**R15**). Using who.is and whois.com, all domains shared the same name servers even if they had different registrars – 162.159.24.201/ns1.dns-parking.com/ns2.dns-parking.com (**R16**). The technique identified consists of phishers using subdomain tricks, namely redirecting to compromised sites with custom subdomains for evasion (**T9**). If attackers use different evasion techniques, then OSINT analysis should be more comprehensive by including the tactic of checking subdomain names as domain names using who.is, whois.com tools (**A8**). In this way, the technique of mixing valid domain names can be used to obtain a subdomain name (**T10**). The findings led to other compromised websites from China, Spain, Pakistan, and the UAE (**R17**). This information led to a new technique that combined many domain names as subdomains for evasion and confusion.

In the next step of tracking digital identity, we carried out cross-social platform checking on Facebook (**A9**) and found that the website theteachingmentors.com was associated with the Teaching Mentors Facebook page (**R18**). The dialling code mentioned on this page led to Pakistan (**R19**). This finding confirmed the technique of using compromised identity for legitimacy (**T11**).

Finally, we also checked for any scam or malicious disclosed activity (**P7**) by verifying all the websites in the virustotal.com, scamadviser.com and webparanoid.com databases (**A10**). The Virus Total results for quoxc.moneyflowgroup.com revealed one security vendor flagging this URL as malicious (**R20**). No other security vendors flagged these websites for malicious activity, such as scam or phishing campaigns (**T12**).

To conclude, all seven purposes of the analysis involved 10 actions that had 20 results and revealed 12 techniques used in this case of cognitive hacking that blended information operations with cyber threat capabilities.

3.5. Tagging Research with the DISARM Framework

The next step in proving a malicious campaign is to match the technological determinants of cognitive hacking with the patterns of influence operations. In this case, we labelled our technical findings under the DISARM framework using DISARM Word Plug-In. Finding attacker behaviours and identifying their tactics and techniques create a behavioural profile based on the DISARM Red Framework – incident creator TTPs, which was useful for determining kill chain attacks.

The use of anonymous and generic names on social platforms (**T1**) is associated with Create Inauthentic Social Media Pages and Groups [T0007], Identify Social and Technical Vulnerabilities: Identify Media System Vulnerabilities [T0081.008]), Create Personas [T0097], Conceal Information Assets: Use Pseudonyms [T0128.001], Conceal Information Assets: Conceal Network Identity [T0128.002], Create Inauthentic Accounts [T0090], Create Inauthentic Accounts: Create Anonymous Accounts [T0090.001], and Conceal Information Assets: Use Pseudonyms [T0128.001].

The compromise of the public newsfeed website to plant fake article on a public newsfeed that mixes the false information with the excerpts copied from an authentic news website (**T2**) is associated with the Compromise Legitimate Accounts [T0011], Compromise Legitimate Accounts [T0011], Distort Facts [T0023], Distort Facts: Edit Open-Source Content [T0023.002], Flooding the Information Space: Bots Amplify via Automated Forwarding and Reposting [T0049.003], Reuse Existing Content [T0084], Reuse Existing Content: Use Copypasta [T0084.001], and Reuse Existing Content: Plagiarize Content [T0084.002].

AI-generated voice-over for a stock video (**T3**) and the use of deepfake content (**T4**) are mentioned in Create Clickbait [T0016], Develop Image-Based Content: Develop AI-Generated Images (Deepfakes) [T0086.002], Develop Video-Based Content: Develop AI-Generated Videos (Deepfakes) [T0087.001], and Develop Audio-Based Content: Develop AI-Generated Audio (Deepfakes) [T0088.001].

The use of a translated fake privacy policy (**T6**) machine is identified in Distort Facts: Edit Open-Source Content [T0023.002], Reuse Existing

Content [T0084], Reuse Existing Content: Use Copypasta [T0084.001], Reuse Existing Content: Plagiarise Content [T0084.002], and Reuse Existing Content: Deceptively Labeled or Translated [T0084.003].

Redirecting to comprised or captured websites (**T8**) and using compromised identities for legitimacy (**T11**) are associated with compromise legacy accounts [T0011], build networks: create organisations [T0092.001], prepare assets impersonating legitimate entities [T0099], control information environments through intensive cyberspace operations: conduct server redirect [T0123.004], conventional operational activity: Redirect URLs [T0129.008], and create automatic websites [T0013].

Deleting tracking records (**T7**), customising subdomains (**T9**), mixing valid domain names to obtain a subdomain name (**T10**), or hiding a video in a shuffle playlist **(T5**) are not tagged as evasion techniques in DISARM, but these techniques are correlated with Compromise Legitimate Accounts [T0011], Harass: Threaten to Dox [T0048.003], Harass: Dox [T0048.004], Map Target Audience Information Environment [T0080], Identify Social and Technical Vulnerabilities [T0081], Infiltrate Existing Networks [T0094], and Conceal Information Assets: Launder Information Assets [T0128.004].

The malicious activity of the website identified as scam or phishing campaigns (**T12**) is associated with the Control Information Environment through Offensive Cyberspace Operations [T0123] and Make Money: Scam [T0137.002].

Summarising the TTPs uncovered by the OSINT analysis based on the PART model and tagging them under the DISARM framework, the overall picture of this cognitive hacking case revealed the intentions, persistence, and level of sophistication of the influencing actors behind this misleading campaign.

4. Discussion

The practice of using the DISARM framework for analysing the cognitive hacking case in Romania was proved to be as reliable as the analysis of the targeted misinformation, disinformation, and malinformation (MDM) campaigns driven by two specific Russian campaigns in Italy surrounding the war in Ukraine [57].

This level of analysis has limitations in terms of technical attribution. There was evidence of hostile actors, such as Russian privacy policy pages, but nothing to conclusively tie it to a specific hostile

state. The use of national symbols or remarks about the state's reputation could be much more related to a FIMI operation. Without them, these misleading ads could easily be associated with common cybercrime.

At the same time, as **R10** has proven, attackers can cover their digital fingerprints by changing the name and activity of the channels used in the influence campaign, which makes their tracking more difficult. Given the absence of a clear public beneficiary for this advertising campaign, the following question arises: Who would invest funds to establish this content engine and execute the malvertising campaign? Perhaps the tech companies managing the advertising platforms could easily uncover the answer by tracking the source of funding. Instead, as researchers, you have to hope for a potential answer by tracking operational patterns over time, leveraging the identified modus operandi based on TTPs.

Another concern relates to the possible impact of these misleading advertising campaigns whose efforts to materialise do not seem to make sense at first sight. A high level of uncertainty, no obvious financial motivation, and the absence of any legally responsible entity could be the predictors of the cognitive hacking deployed for social harm or political pressure.

In the context of countering cyber-enabled FIMI, any practical approach to existing tools can improve defence strategies by updating TTPs, similar to the sharing databases used in cybersecurity. When confronted with a hybrid threat, response actions should be combined starting at the strategic level. In our research, the meta-analysis based on the strategic structured PART model could be replicated and improved by other researchers, building on other cases and with more sophisticated tools. Tagging the findings into the DISARM framework can prove their two-fold utility. On the one hand, it can confirm the effectiveness of the framework by linking it to already identified techniques; on the other hand, it can improve the framework by adding new techniques, given the constant evolution of cyber threats.

5. Conclusions

This in-depth analysis of a cognitive hacking case can provide the basis for a set of new methodologies for exposing malicious interference in people's minds. Starting from nothing more than the identification of apparently irrelevant video ads, which are usually ignored by analysts, using open tools, and accessing public

databases can reveal a malicious scheme that blends information operations with cyber threat capabilities. This analysis was carried out from the perspective of two regular Internet users with an average level of digital literacy and awareness of cyber threats and no sophisticated online tools.

The research involved 10 steps, provided 20 results and revealed 12 techniques used in this case of cognitive hacking. Using AI-generated content in deepfake ads, hijacking websites and planting fabricated content under anonymity, abusing social networks, and purchasing targeted advertisements to manipulate vulnerable social groups are well-known tactics and techniques used in the preparedness stage of cyber influence operations.

The remaining question is, what to do with such cases that, at first sight, appear to have no discernible association with any specific entity or purpose or that do not overtly indicate any explicit threat. Should they be dismissed as irrelevant? Based on this in-depth analysis, when there is no clear evidence of what entity is involved and for what purpose, such situations can provide early warning of a potential attack in the preparatory stage. Detecting these signs early, before the actions as such manage to alter the analysts' perspective on what happens and how it happens, can reinforce defence mechanisms, and thwart malicious actions in their infancy, which is the most desirable scenario for defence. The extensive analysis of the identified case builds confidence that applying the DISARM framework to cyber-enabled influence campaigns can be useful for anticipating cyfluence and FIMI operations, even when such operations do not appear to have specific, immediately identifiable perpetrators or purposes.

References

[1] A.K. Sood, R.J. Enbody, "Malvertising – exploiting web advertising," *Computer Fraud & Security*, vol. 2011, no. 4, pp. 11–16, 2011, doi: 10.1016/S1361-3723(11)70041-0.

[2] C. Dwyer, A.A. Kanguri, "Malvertising – A rising threat to the online ecosystem," *Journal of Information Systems Applied Research (JISAR)*, vol. 10, p. 29, 2017.

[3] M. Willett, "The cyber dimension of the Russia–Ukraine war," *Survival (Lond)*, vol. 64, no. 5, pp. 7–26, 2022, doi: 10.1080/00396338.2022.2126193.

[4] K. Poireault. (Nov. 18, 2023). "Russian APT Sandworm disrupted power in Ukraine using novel OT techniques," *Infosecurity Magazine*. [Online]. Available: https://www.infosecurity-magazine.com/news/russia-sandworm-disrupted-power. [Accessed: Jun. 25, 2023].

[5] G. Pakharenko, "Cyber operations at Maidan: A first-hand account," in *Cyber War in Perspective: Russian Aggression against Ukraine*, K. Geers, Ed. Tallinn: NATO–CCD-COE Publications, 2015, pp. 59-66. [Online]. Available: https://ccdcoe.org/uploads/2018/10/Ch07_CyberWarinPerspective_Pakharenko.pdf [Accessed: Jun. 25, 2023].

[6] J. Przetacznik, S. Tarpova. (2022). *Russia's war on Ukraine: Timeline of cyber-attacks*. [Online]. Available: https://epthinktank.eu/2022/06/21/russias-war-on-ukraine-timeline-of-cyber-attacks. [Accessed: Jun. 25, 2023].

[7] Microsoft. (2022). *Defending Ukraine: Early lessons from the cyber war*. [Online]. Available: https://blogs.microsoft.com/on-the-issues/2022/06/22/defending-ukraine-early-lessons-from-the-cyber-war. [Accessed: Apr. 11, 2023].

[8] G. Wilde. (Dec. 12, 2022). *Cyber operations in Ukraine: Russia's unmet expectations*. [Online]. Available: https://carnegieendowment.org/research/2022/12/cyber-operations-in-ukraine-russias-unmet-expectations?lang=en. [Accessed: Mar. 9, 2023].

[9] A. MacDonald, R. Ratcliffe. (Nov. 18, 2023). *Cognitive warfare: maneuvering in the human dimension*. [Online]. Available: https://www.usni.org/magazines/proceedings/2023/April/cognitive-warfare-maneuvering-human-dimension. [Accessed: Dec. 03, 2023].

[10] D. Susser, B. Roessler, H. Nissenbaum, "Online manipulation: Hidden influences in a digital world," *Georgetown Law Technology Review*, vol. 1, pp. 1–45, 2019, doi: 10.2139/ssrn.3306006.

[11] R. Medrano. (2023). *Cognitive warfare: Halting the Russian sphere of influence*. [Online]. Available: https://apps.dtic.mil/sti/trecms/pdf/AD1200752.pdf. [Accessed: Dec. 12, 2023].

[12] Y. Danyk, C.M. Briggs, "Modern cognitive operations and hybrid warfare," *Journal of Strategic Security*, vol. 16, no. 1, pp. 35–50, 2023, doi: 10.2307/48718245.

[13] P. Krawczyk, J. Wiśnicki, "Russia's social-impact operations in the context of cognitive warfare in Ukraine in 2022," *Cybersecurity and Law*, vol. 9, no. 1, pp. 194–203, 2023, doi: 10.35467/cal/169315.

[14] J.F. Tripp, N.K. Lankton, D.H. Mcknight, J. Tripp, "Technology, humanness, and trust: rethinking trust in technology," *Journal of the Association for Information Research*, vol. 16, no. 10, pp. 880–918, 2015, doi: 10.17705/1jais.00411.

[15] D.B. Hollis, "The influence of war; the war for influence," *Temple International & Comparative Law Journal*, vol. 32, no. 1, pp. 31, 2018.

[16] G. Cybenko, A. Giani, P. Thompson, "Cognitive hacking," *Advances in Computers*, vol. 60, pp. 35–73, 2004, doi: 10.1016/S0065-2458(03)60002-1.

[17] NATO Allied Command Transformation. (Nov. 18, 2023). *Cognitive warfare: Strengthening and defending the mind*. [Online]. Available: https://www.act.nato.int/article/cognitive-warfare-strengthening-and-defending-the-mind. [Accessed: Dec. 13, 2023].

[18] A. Ertan, K.H. Floyd, P. Pernik, T. Stevens. (2020). *Cyber threats and NATO 2030: Horizon scanning and analysis*. [Online]. Available: https://ccdcoe.org/library/publications/cyber-threats-and-nato-2030-horizon-scanning-and-analysis. [Accessed: Jun. 25, 2023].

[19] O. Backes, A. Swab. (2019). *Cognitive warfare: The Russian threat to election integrity in the Baltic states*. [Online]. Available: https://www.belfercenter.org/publication/cognitive-warfare-russian-threat-election-integrity-baltic-states. [Accessed: Dec. 23, 2023].

[20] A. Bernal, C. Carter, I. Singh, K. Cao, O. Madreperla. (2020). *Cognitive warfare: An attack on truth and thought*. [Online]. Available: https://www.innovationhub-act.org/sites/default/files/2021-03/Cognitive%20Warfare.pdf. [Accessed: Nov. 18, 2023].

[21] N. Beauchamp-Mustafaga, "Cognitive domain operations: The PLA's new holistic concept for influence operations," *China Brief*, vol. 19, no. 16, 2019. [Online]. Available: https://jamestown.org/program/cognitive-domain-operations-the-plas-new-holistic-concept-for-influence-operations. [Accessed: Nov. 18, 2023].

[22] L. Hauser, *Coordinated chaos: Synchronized cyberwarfare and disinformation attacks.* The Project on International Peace and Security. Williamsburg, VA: Global Research Institute, 2022.

[23] R. Burda, *Cognitive warfare as part of society never-ending battle for minds*. *Information-based behavioral influencing and Western practice paper series*. The Hague: The Hague Centre for Strategic Studies, Jun 23, 2023.

[24] M. Alazab, "Russia is using an onslaught of cyber attacks to undermine Ukraine's defense capabilities," *The Conversation*, 24 Feb. 2022. [Online]. Available: https://theconversation.com/russia-is-using-an-onslaught-of-cyber-attacks-to-undermine-ukraines-defence-capabilities-177638. [Accessed: Mar. 30, 2022].

[25] A. Wahlstrom, A. Revelli, S. Riddell, D. Mainor, R. Serabian. (2022). *The IO offensive: Information operations surrounding the Russian invasion of Ukraine*. [Online]. Available: https://www.mandiant.com/resources/blog/information-operations-surrounding-ukraine. [Accessed: Nov. 18, 2023].

[26] I. Yonat. (2023). *Hostile influence campaigns, cyber security and AI*. [Online]. Available: https://www.linkedin.com/posts/itai-y-5731a146_hostile-influence-campaigns-cyber-security-activity-7077365346583609344-44PL. [Accessed: Nov. 18, 2023].

[27] D. Pereira. (2023). *Cognitive infrastructure worldwide is under attack in "the worst cognitive warfare conditions since WWII."* [Online]. Available: https://www.oodaloop.com/archive/2023/11/08/cognitive-infrastructure-worldwide-is-underattack-in-the-worst-cognitive-warfare-conditions-since-wwii. [Accessed: Nov. 18, 2023].

[28] B. Gourley. (2019). *America's most critical infrastructure is also our most neglected infrastructure*. [Online]. Available: https://www.oodaloop.com/archive/2019/09/03/americas-most-critical-infrastructure-is-also-our-most-neglected-infrastructure. [Accessed: Nov. 18, 2023].

[29] Swedish Psychological Defense Agency. (2023). [Online]. Available: https://www.mpf.se/en/about-us. [Accessed: Nov. 18, 2023].

[30] Joint Research Centre EU. (2020). *Cybersecurity, our digital anchor: A European perspective*, Publications Office of the European Union, Luxembourg. [Online]. Available: https://publications.jrc.ec.europa.eu/repository/handle/JRC121051. [Accessed: Nov. 18, 2023].

[31] A. Polyakova, S.P. Boyer, B.-R. Bosch. (2018). *The future of political warfare: Russia, the West, and the coming age of global digital competition*. [Online].

Available: https://www.brookings.edu/articles/the-future-of-political-warfare-russia-the-west-and-the-coming-age-of-global-digital-competition. [Accessed Mar. 30, 2022].

[32] P.N. Petratos, "Misinformation, disinformation, and fake news: Cyber risks to business," *Business Horizons*, vol. 64, no. 6, pp. 763–774, 2021, doi: 10.1016/j.bushor.2021.07.012.

[33] US Department of State. (2020). *GEC special report: Pillars of Russia's disinformation and propaganda ecosystem*. [Online]. Available: https://www.state.gov/russias-pillars-of-disinformation-and-propaganda-report. [Accessed: Mar. 30, 2022].

[34] M. Baezner, *Cyber and information warfare in the Ukrainian conflict*. ETH Zürich: Center for Security Studies (CSS), 2018.

[35] P. Brangetto, M. A. Veenendaal. (2016). "Influence cyber operations: The use of cyberattacks in support of influence operations." 8th International Conference on Cyber Conflict, 2016, pp. 113–126. [Online]. Available: https://ccdcoe.org/uploads/2018/10/Art-08-Influence-Cyber-Operations-The-Use-of-Cyberattacks-in-Support-of-Influence-Operations.pdf. [Accessed: Mar. 30, 2022].

[36] European Union Agency for Cybersecurity – ENISA. (2021). *ENISA threat landscape 2021*. [Online]. Available: https://www.enisa.europa.eu/publications/enisa-threat-landscape-2021. [Accessed: Mar. 30, 2022].

[37] European Commission. (Dec. 14, 2020). *The EU's cybersecurity strategy for the digital decade*. [Online]. Available: https://digital-strategy.ec.europa.eu/en/library/eus-cybersecurity-strategy-digital-decade-0. [Accessed: Nov. 18, 2023].

[38] Y. Danyk, T. Maliarchuk, C. Briggs, "Hybrid war: High-tech, information and cyber conflicts," *Connections: The Quarterly Journal*, vol. 16, no. 2, pp. 5–24, 2017, doi: 10.11610/connections.16.2.01.

[39] D. Ardia, E. Ringel, V. S. Ekstrand, A. Fox, "Addressing the decline of local news, rise of platforms, and spread of mis-and disinformation online: A summary of current research and policy proposals," *UNC Legal Studies Research Paper,* 15 Jan. 2021. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3765576. [Accessed: Nov. 18, 2023].

[40] B. Martens, L. Aguiar, E. Gomez-Herrera, F. Mueller-Langer. (2018). *The digital transformation of news media and the rise of disinformation and fake news – An economic perspective*. [Online]. Available: https://ec.europa.eu/jrc. [Accessed: Nov. 18, 2023].

[41] J.T. Rob, J.N. Shapiro. (Jun. 12, 2022). *A brief history of online influence operations*. [Online]. Available: https://www.lawfareblog.com/brief-history-online-influence-operations. [Accessed: Nov. 18, 2023].

[42] P. M. Duggan, "Strategic development of special warfare in cyberspace," *Joint Force Quarterly 79*, vol. 79, no. 4, pp. 46–53, 2015.

[43] C. Whyte, A. T. Thrall, B. M. Mazanec, Eds., *Information warfare in the age of cyber conflict*. London, UK, New York, NY, USA: Routledge Taylor & Francis Group, 2021.

[44] C. Whyte, "Cyber conflict or democracy 'hacked'?" How cyber operations enhance information warfare," *Journal of Cybersecurity,* vol. 6, no. 1, 2020, pp. 1-17, doi: 10.1093/cybsec/tyaa013.

[45] R. Manwaring, J. Holloway, "Resilience to cyber-enabled foreign interference: Citizen understanding and threat perceptions," *Defense Studies*, vol. 23, no. 2, pp. 334–357, 2022, doi: 10.1080/14702436.2022.2138349.

[46] M.A. Gomez, "Cyber-enabled information warfare and influence operations," in *Information warfare in the age of cyber conflict*, C. Whyte, A.T. Thrall, B.M. Mazanec, Eds., London: Routledge Taylor & Francis, 2021, pp. 132–146.

[47] M. Wigell, "Hybrid interference as a wedge strategy: A theory of external interference in liberal democracy," *International Affairs*, vol. 95, no. 2, pp. 255–275, 2019, doi: 10.1093/ia/iiz018.

[48] M. Weissmann, N. Nilsson, B. Palmertz, P. Thunholm, Eds., *Hybrid warfare: Security and asymmetric conflict in international relations.* London: I.B. Tauris, 2021.

[49] M.E. Zurko, "Disinformation and reflections from usable security," *IEEE Security and Privacy*, vol. 20, no. 3., pp. 4–7, 2022, doi: 10.1109/MSEC.2022.3159405.

[50] I. Lella, M. Theocharidou, E. Tsekmezoglou, R. Svetozarov Naydenov, C. Ciobanu, A. Malatras. (2022). *ENISA threat landscape 2022*. [Online]. Available: https://www.enisa.europa.eu/publications/enisa-threat-landscape-2022 [Accessed: Aug. 20, 2023].

[51] C. R. Walker, S.-J. Terp, P. C. Breuer, C. L. Crooks, "Misinfosec," Companion Proceedings of The 2019 World Wide Web Conference. Association for Computing Machinery, May 13, pp. 1026–1032, 2019. doi: 10.1145/3308560.3316742.

[52] DISARM Foundation. (Nov. 12, 2023). *DISARM framework*. [Online]. Available: https://www.disarm.foundation/framework. [Accessed: Nov. 18, 2023].

[53] EU-US Trade and Technology Council. (May 31, 2023). *Trade and Technology Council Fourth Ministerial – Annex on foreign information manipulation and interference in third countries*. [Online]. Available: https://www.eeas.europa.eu/eeas/trade-and-technology-council-fourth-ministerial-%E2%80%93-annex-foreign-information-manipulation-and_en. [Accessed: Mar. 22, 2024].

[54] S. Terp, P. Breuer. (2022). "DISARM: A framework for analysis of disinformation campaigns." 2022 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA), Salerno, Italy, pp. 1–8, doi: 10.1109/CogSIMA54611.2022.9830669.

[55] H. Newman. (2022). *Foreign information manipulation and interference defense standards: Test for rapid adoption of the common language and framework "DISARM."* [Online]. Available: https://stratcomcoe.org/publications/foreign-information-manipulation-and-interference-defence-standards-test-for-rapid-adoption-of-the-common-language-and-framework-disarm-prepared-in-cooperation-with-hybrid-coe/253. [Accessed: Mar. 30, 2022].

[56] R. Arcos. (2018). *Post event analysis of the hybrid threat security environment: assessment of influence communication operations.* [Online]. Available: https://www.hybridcoe.fi/publications/hybrid-coe-strategic-analysis-12-postevent-analysis-of-the-hybrid-threat-security-environment-assessment-of-influence-communication-operations. [Accessed: Nov. 12, 2023].

[57] M. Lesser, H.J. Stern, S.J. Terp. (2022). "Countering Russian misinformation, disinformation, malinformation and influence campaigns in Italy surrounding the Russian invasion of Ukraine," *International Forum on Digital and Democracy 2022*. [Online]. Available: https://ceur-ws.org/Vol-3289/paper2.pdf. [Accessed: Nov. 12, 2023].

ANNEX 1. Screenshots of the case identification stage

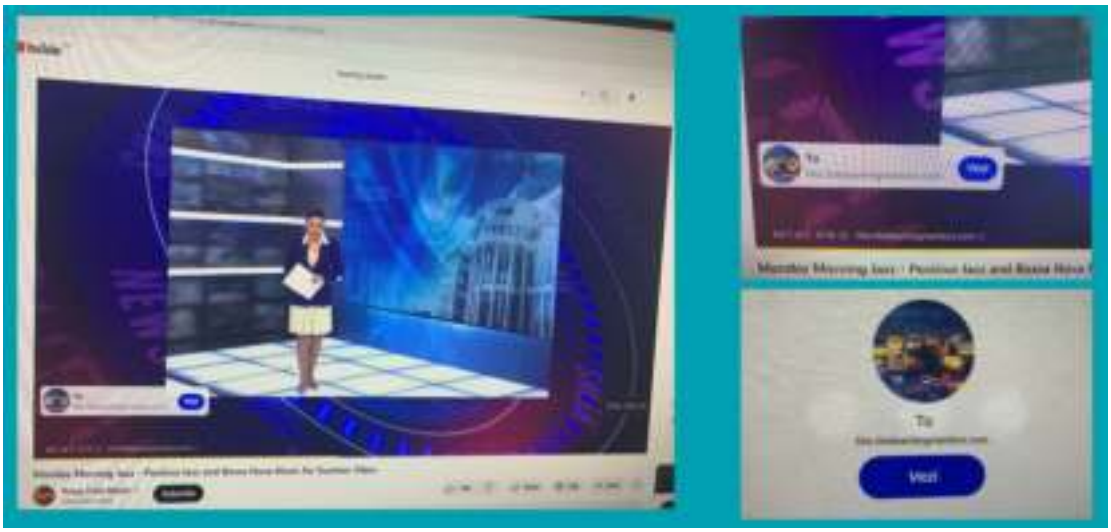
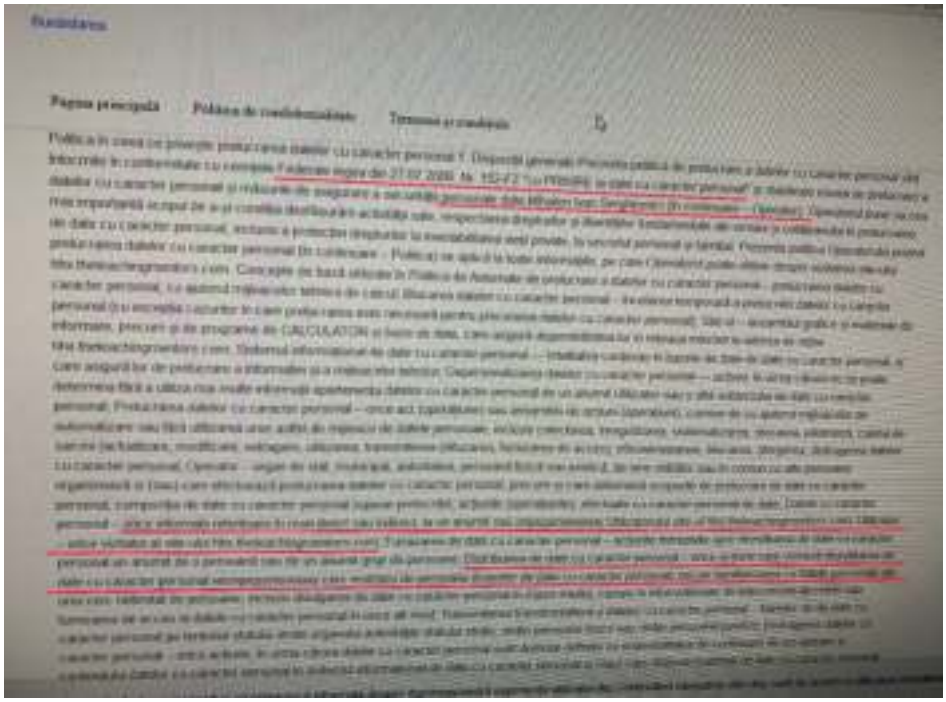


Photo captured from the first video ad promoted on YouTube.

Photo captured from the second video ad promoted on YouTube.

ANNEX 2. Screenshots of digital analysis

Screenshot of the website gute.mycalculat.com



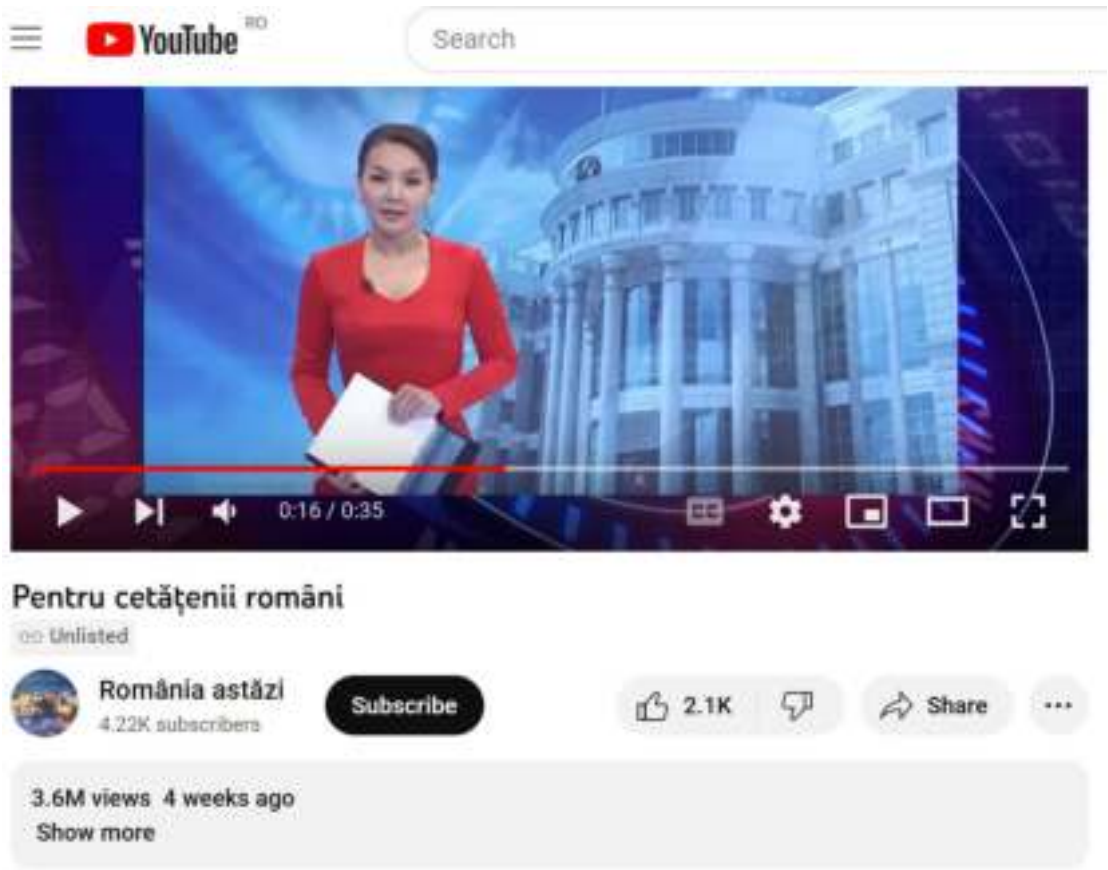
Capture of the fake policy.

ANNEX 3. Screenshots of OSINT results

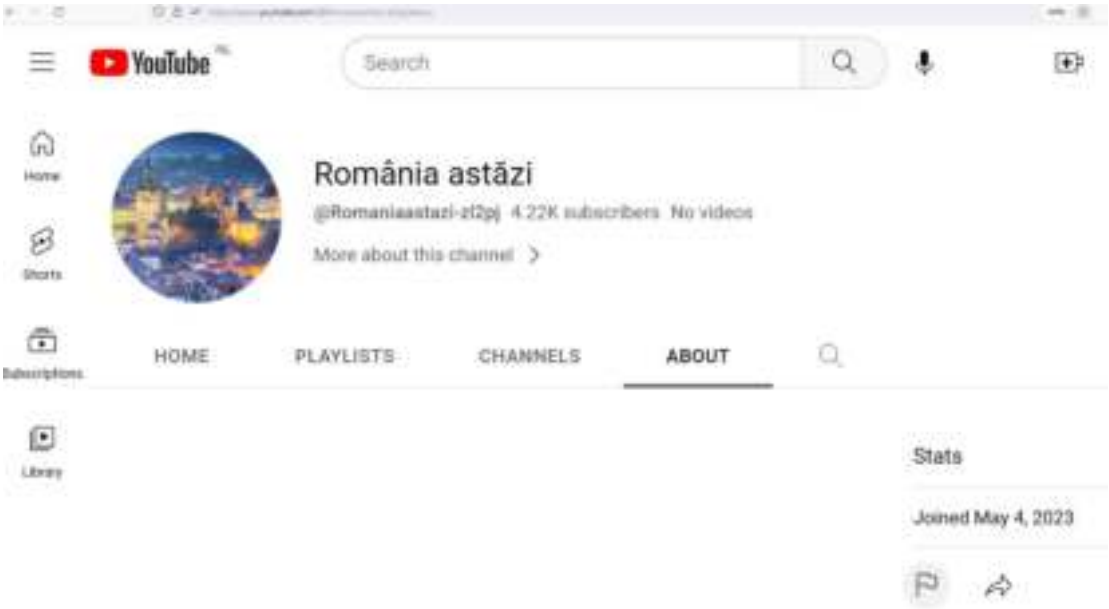
R1: The screenshot posted by a Reddit user who uncovered the original YouTube channel managing the video ad campaign.

R2: The fake article planted on weeklynewsfeed.com revealed by another Reddit user.

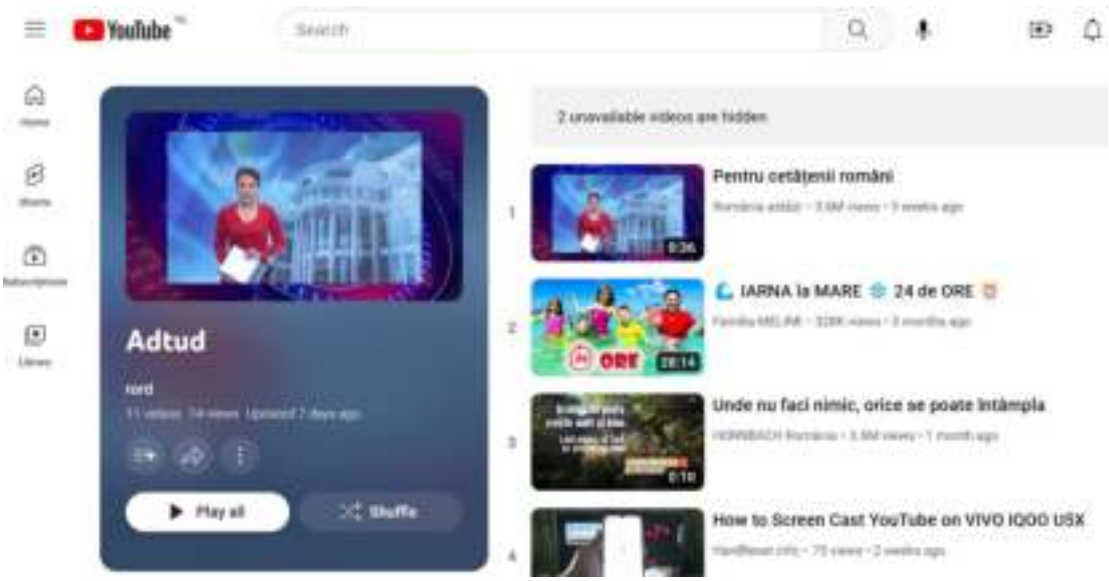
R4: Google images match the image footage with the Ak Orda Presidential Palace in Kazakhstan.



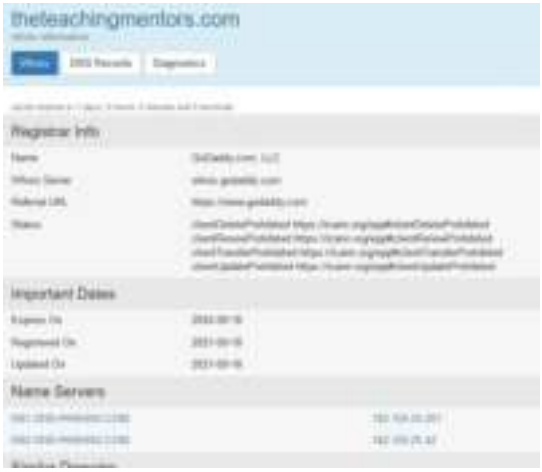
R6: The video ad named '*Pentru cetățenii români*' ('For Romanian Citizens') was posted on 6 May 2023 by the *România astăzi* (România Today) channel and reached over 3.65 million viewers and received 2.1K like reactions and 23 comments in 1 month (figures as on 22 May 2023).

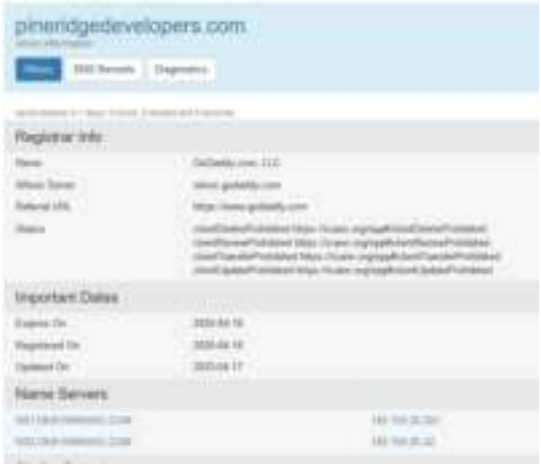


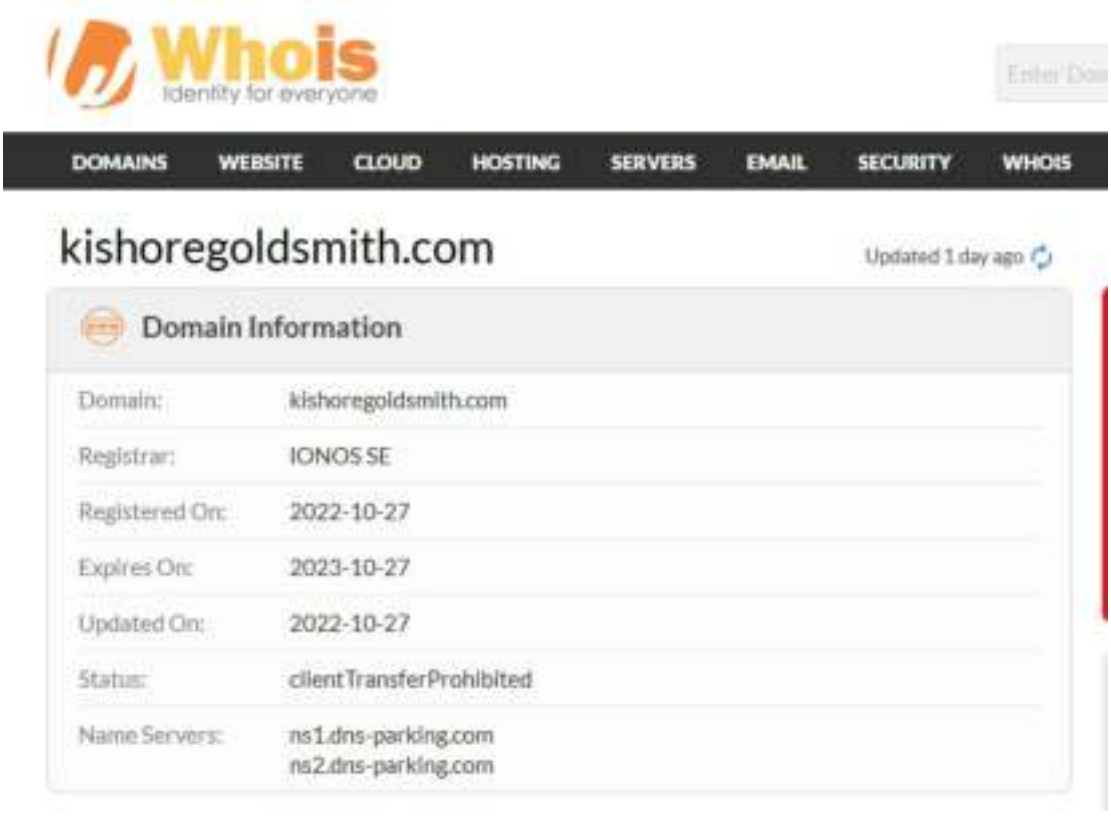
R7: The YouTube channel ID @Romaniaastazi-zl2pj has 4.22K subscribers, joining YouTube on 4 May 2023.



R8: The video ad named 'For Romanian Citizens' identified in the playlist Adtud of the user *rord* aka @peisaj131 – https://www.youtube.com/@peisaj131







R14: All identified websites have the same name servers in the same class C subnet (the first three numbers of their IPs are identical). **R16:** All domains share the same name servers even if they have different registrars – 162.159.24.201/ns1.dns-parking.com/ns2.dns-parking.com