Comparative Analysis of Disinformation Regulations: A Preliminary Analysis

Antonella Calò(✉), Antonella Longo, and Marco Zappatore

Department of Engineering for Innovation, University of Salento, 73100 Lecce, Italy
{antonella.calo,antonella.longo,
marcosalvatore.zappatore}@unisalento.it

Abstract. The spread of Fake News poses a significant threat to democracy and public discourse. Instances of disinformation have had serious consequences, such as undermining election integrity and reducing vaccine trust over the recent years. Moreover, terminological variations and digital neologisms hinder consensus among scholars. The Internet proliferation has intensified the challenge of managing false content, demanding technological tools and regulations. This study explores the complexities of Fake News and the need for regulatory measures. A comparative law methodology is used to analyze international and European regulations concerning Fake News. Social media policies and reporting methods are also investigated, aiming to find ways to combat misinformation effectively. On the one hand, a fragmented regulatory framework both at a European and International level is revealed. To cope with this scenario, a multilingual ontology to harmonize definitions and facilitate compliance is proposed. On the other hand, the crucial role of Social Media policies, their algorithms' transparency, and educating roles are considered. This leads to the need for an enhanced regulation of social media, educational initiatives of digital media literacy, and AI-driven news apps to provide trusted sources and manage misinformation in a better way.

Keywords: Disinformation · Fake News Regulation · Digital Media Literacy · International Law · European Law · Social Media Policies · Reporting Systems

1 The Threat of Fake News: Growing Challenges and the Need for Regulation

The plurality of communication channels and the spread of Fake News are widespread phenomena today [1]. The more this phenomenon increases, the more it represents a threat to democracy and to the public sphere. Indeed, there is a significant number of examples of disinformation, and Fake News that caused serious issues, such as the most famous one during Donald Trump's 2016 elections, according to which Russian hackers were attacking the vote-system software to track and manipulate election results [2] or the 600% increase of disinformation about vaccines after the decision of the Court of Rimini to recognize that MMT (Measles, Mumps, and Rubella) vaccine caused autism in a child. This caused low trust in the health systems and vaccines, thus leading to a reduction in child immunization [3]. Finally, the huge infodemic during the COVID-19 pandemic was another noteworthy source of disinformation.

T. P. Sales et al. (Eds.): ER 2023 Workshops, LNCS 14319, pp. 162–171, 2023.
https://doi.org/10.1007/978-3-031-47112-4_15

The terms related to Fake News have different definitions, and the phenomena assume different forms, various authors, and multiple motives, such as commercial sensation contents, ideologically driven news, or state sponsored misinformation to influence users [4]. The phenomenon of misinformation/disinformation has always existed, but with the rise of the Internet, it has now increased significantly, causing a more difficult management of false contents. This led to the need of technological tools capable of detecting Fake News, and to the creation of regulations and measures. The latter is the topic of the present paper.

The research questions this work will address are:

1. Which are the regulations at an international and European Level related to the dissemination of Fake News?
2. How to ensure compliance, and harmonization among those regulations also from a terminological standpoint?
3. Is it possible to report Fake News on social media? And how can social networks educate their users?

The paper is organized as follows: the second section will focus on the state of the art and legal/regulatory framework. In the third section we will provide the methodology followed to check the possibility to report Fake News on social media and the application of the methodology. The fourth section will present the results. Finally, conclusions and the future work will be drawn.

2 State of the Art: Legal, and Regulatory Framework in Public/Private Sectors

The study of legal literature concerning misinformation, disinformation, and Fake News reveals a fundamental challenge rooted in the use of terms and language. The evolution of these terms, particularly the adoption of "Fake News" as a label for what was once called "false information," has not brought consensus among scholars and experts [5]. Claire Wardle and Hossein Derakhshan call these challenges "Information disorder" [6]. They showed that the term "Fake News" has been used to refer to different phenomena and results inadequate, therefore the authors proposed three notions:

- **Dis-information**: Information that is false and deliberately created to harm a person, social group, organization, or country.
- **Mis-information:** Information that is false, but not created with the intention of causing harm.
- **Mal-information**: Information that is based on reality, used to inflict harm on a person, organization, or country.

Different interpretations of these terms have been described in [7], while the European definitions can be checked in [8].

The phenomenon of Fake News in the current "post-truth" era is multifaceted and complex, and it often exploits emotions to polarize readers and elicit strong reactions. However, its implications go beyond emotional manipulation. Fake News frequently intertwines with criminal behaviors, such as hate speech and revenge porn facilitated by technologies like DeepFake and DeepNudes. Semantic challenges arise when addressing cybercrimes and developing regulations, particularly regarding terminology and the discrepancies among different legal systems. Supranational cybercrime cases that transcend national jurisdictions pose additional difficulties, as legislative differences hinder comprehension and interoperability [9]. Fake News, and all the phenomenon of disinformation, and misinformation can be considered an example of cybercrime; and, because of the Internet, they certainly circulate out of national jurisdictions. Cybercrimes, and these types of contents shared online, share new abstract terminologies, or neologisms such as: "DeepFake"; "eco-chambers"; "Clickbait", and so on. The fragmented terminology, and their continuous development also cause problems in translation. Furthermore, because of the lack of digitalization, and the lack of trust in digital tools, Artificial Intelligence and Machine Learning, legal professional translators refuse to use the so-called CAT (Computer-Aided translation Tools) tools, leading to a low interoperability between the legal terms [10], and mistranslations of regulations, and lack of compliance. Some examples of mistranslations of the European GDPR regulation in the Italian, Icelandic, Finnish, and Slovak languages can be found in [10].

This is why, regulations appear fragmented, and ontologies have been created to face these problems [9, 11, 12]. In the case of revenge porn, or other types of abstract concepts such as in [11, 13], different European Member States have tried to face the issue with different regulations, which yet do not seem to be harmonized at European level. The situation is even more fragmented when it comes to regulating Fake News. Indeed, the lack of proper definitions, and interpretations of the terms, causes issues when regulating these phenomena, resulting in a lack of regulations (or not harmonized ones). Furthermore, concerns about public and private censorship risk violating Article 19 of the United Nations' right to freedom of speech and expression, as well as other directives and regulations safeguarding freedom of expression [5, 14].

To mitigate these risks and promote effective regulation, collaboration with social media platforms has become crucial. Social media algorithms, designed to recognize users' preferences, often reinforce confirmation bias, and create echo chambers. The opacity of algorithms, often referred to as "black boxes", contributes to the erosion of trust in technology and data sharing practices. The adoption of Explainable AI (XAI) can enhance transparency and address biases by providing insights into the involvement of personal data [15]. Research on XAI in societal challenges has shown promising results in enhancing transparency, mitigating biases, and combating disinformation [16].

In the subsequent section, we will delve into a comparative analysis of existing regulations at both public and private levels, examining the strategies employed by social media platforms to report and combat disinformation. By exploring these approaches, we aim to shed light on the complex landscape of Fake News regulation and identify potential areas for improvement and collaboration.

3 Case Studies: Comparative Methodology Analysis of Regulations and Media Policies

We will proceed with a comparative law methodology [17], which will give us the possibility to understand whether regulations are present or not, if European directives have been implemented or not, if they are compliant among them, and finally if they have been criticized in a positive or negative way. This will allow us to have a better understanding of the next steps to take to better face the disinformation/Fake News phenomena. We will present a multilingual ontology that when larger and more modular will support the drafting and translation of regulations processes.

A comparative method will be also used to analyze some of the social media policies, and their reporting methods to better understand if, and how to cooperate/regulate them.

3.1 Case Study 1: Comparison of Public Regulations at the International and European Level

A comparative law methodology is employed to analyze public international, European, and national directives, regulations, and proposals regarding disinformation [6, 18–20]. The first European Union attempt to combat disinformation involved the Commission Communication in 2017"Tackling Disinformation: A European approach," which was a non-binding instrument aimed at fostering cooperation with social media platforms. This led to the autoregulation soft law "**Code of Conduct**", which has the goal to stop ads of accounts that disseminate disinformation, improve transparency, cooperate with Social Media in deleting fake accounts/bots, improve official sources with trusted news, and allow academics to access data. However, this approach was criticized for potentially granting excessive power to social media platforms and risking private censorship. Germany and France attempted to implement the code of conduct but faced opposition for potentially infringing upon freedom of speech. The German implentation (**NetzDG**) of 2017, after defining the meaning of "lawful" and "manifestly lawful", makes the social media with more than 200 reports to share a public report on how they faced the issues; added an administrative authority of federal justice; and poses penal sanctions up to 20 millions of euros.

The French implementation **(Loi relative à la lutte contre la manipulation de l'information**) of 2018, was created specifically for electoral periods, and added the obligation of transparency. Indeed, platforms had to create reporting systems, and reported Fake News would have been analyzed within 48 h risking sanctions up to 75.000 euros [20]. These implementations and proposals were rejected because they were seen particularly stringent and raised concerns about the right of freedom of speech. Similarly, Italian proposals (such as **DDL Gambaro** which proposed to add fines; **DDL Zanda-Filippin** that followed the German Model, and **Boldrini's Campaign** which focuses on a campaign of Digital Media Literacy) faced rejection for similar reasons, and the current articles 656-bis/ter still reference disinformation using terminology from 1936 [19]. In 2020, the Code of Conduct was implemented with the **Digital Service Act**, a co-regulation instrument that faced illegal and systemic risks; invite platforms to create conduct codes; empower users with more transparency, and create new independent authorities. However, its adoption remains limited, with only France and Germany having implemented it thus far. Germany, implemented it in 2020 with the "**Mstv**" which regulates algorithms, adds labels of social bots, and improve the findability of public service and journalists' due diligence for social media.

France implemented it in 2020 with the "**Loi visant à lutter contre les contenus haineux sur Internet**", which took down fake accounts about Covid-19 within one hours.

Internationally, regulations and proposals have been tested, except for China, which holds different attitudes toward freedom of expression, and Australia, which focuses more on media literacy. Indeed, China's regulation called "Cybersecurity Law" added fines and detection for those who do not respect the rules. It allows to share only registered news media articles. It also created a platform called Piyao, which broadcasts only real news, allows user to report contents and has AI algorithms that detect rumors.

During the same years, in Australia a Taskfor for Fake News led to different codes and legislations, which focused on: Social Media Literacy; a multi-agency body that could address to disinformation risks (among others); Promotion of self-regulation and a new advertising campaign during elector periods called "Stop and Consider" to encourage users to check the sources.

Also the USA faced this topic during these years, through different measures ("Honest Act", "US National Defence Authorisation act", "Stigler Committee"). These required social media to keep copies of ads, the federal governement to deal with propaganda and disinformation, proposed a Digital Agency, and implement antitrust, data protection laws and media policies. However, most of these regulations have been rejected.

3.2 Case Study 2: Facebook vs Twitter Policies Comparison, and Reporting Methods.

The following is a comparison of the two most important social media related to the news policies: Facebook[1] and Twitter[2].

During the last years, Facebook has continuously implemented its policies. The current one, updated to 2023, is the **"Community Standards - integrity and authenticity"** policy, through which the platform removes: physical harm or violence, harmful health misinformation, miracle cures, manipulated media, and election interferences.

In 2022, Twitter (now "X"), through the **"Crisis Misinformation policy"** punished false or misleading information that could harm crisis-affected populations, and decreased the account visibility, adding a warning note to those accounts that violated this policy. Twitter added then two other policies in 2023 **("Synthetic and manipulated media policy" and "Civic Integrity policy"**) which punish those manipulated/false contents which are not satire, memes, animations, and opinions. More specifically, the "civic integrity policy" relates to those contents about elections or civic processes. In both cases, accounts risk to be deleted or labeled.

Both Facebook and Twitter trust on users who can report, and on the work of external fact-checkers services. Facebook relies on fact-checkers, certified through the non-partisan International Fact-checking Network (IFCN)[3].

1 https://transparency.fb.com/en-gb/policies/community-standards/misinformation/

2 https://help.twitter.com/en/rules-and-policies/crisis-misinformation; https://help.twitter.com/en/rules-and-policies/manipulated-media; https://help.twitter.com/en/rules-and-policies/election-integrity-policy.

3 https://transparency.fb.com/en-gb/features/how-fact-checking-works/

4 Discussion

4.1 Coping with Issues in Regulations: An Ontological Proposal (Case Study1)

The European situation regarding the regulation of Fake News is fragmented, as member states have not fully adapted to the Digital Service Act [1]. Terminological inconsistencies persist, with some countries, like Italy, having different interpretations of the term "Fake News" or using outdated terminology. Addressing terminological challenges and regulating internet-related events have been explored in previous works using ontologies [9, 12]. Drawing inspiration from Castano's hierarchy and conceptual models [12], We propose to develop a multilingual ontology that can serve as a reference for the definition and creation of regulations and aid in translation processes. Such an ontology could promote harmonization and compliance in international regulations concerning Fake News.

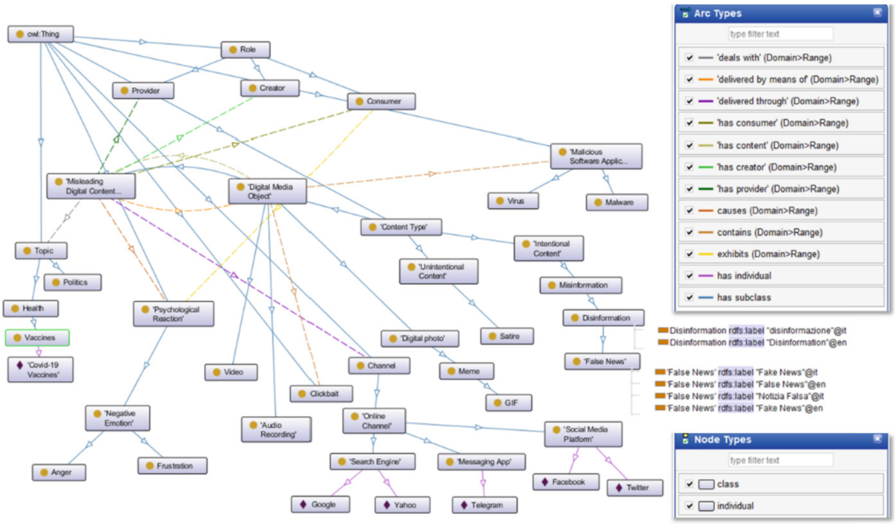

Fig. 1. Multilingual Ontology of terms related to "Misleading Digital Content".

The proposed ontology provides the definitions, translations, and instances of each element. This allows the understanding of a non-unique definition of all elements, in each language, thus eliminating confusion, especially when it comes to the use of "Fake News", "Disinformation", "Misinformation", or neologisms such as "Clickbait". It also provides synonyms for each element in the different languages, modeled as class attributes, and the corresponding semantic relationships, modeled as object properties, showed as arrows in Fig. 1. This ontology can be the starting point for a larger and a modular ontology that would support the regulatory drafting and translation processes.

4.2 Social Campaigns to Regulate Disinformation Through Awareness

At the international level, laws addressing Fake News have faced criticism for potentially infringing upon freedom of speech and granting excessive power to social media platforms. However, positive measures have been observed in the realm of digital literacy and awareness campaigns, which empower individuals to critically evaluate and resist sociocultural manipulation [6, 21]. Educational attainment plays a significant role in enhancing individuals' ability to comprehend and assess information. Younger individuals, especially those with higher levels of education, demonstrate greater awareness and concern for public issues due to their access to reliable information and knowledge [1]. Cybersecurity awareness programs have proven effective in reducing risky online behaviors and fostering a sense of self-responsibility among participants [22]. Consequently, governmental, and private initiatives promoting media literacy, such as "Open the Box[4] "and the "Bad News[5] " game, are promising, although most initiatives primarily target young users. To address this gap, media literacy education through social media should be expanded to reach users of all ages and educational backgrounds. Additionally, promoting information security awareness among smartphone users, particularly concerning data collection methods, is crucial [23].

4.3 Discussing Media Policies and Reporting Methods (Case Study2)

In a world dominated by social media platforms, two key players emerge: Facebook and Twitter, but we will also discuss a new news app called Artifact[6]. These platforms are grappling with the pervasive issue of misinformation and the challenge of maintaining a balance between freedom of speech and responsible content moderation.

Facebook, as we learned in Sect. 4.2, places great importance on respecting national and international regulations. They have implemented robust policies to automatically detect and delete any content that violates community standards or legal regulations. In some countries, Facebook utilizes advanced technologies to identify possible misinformation and promptly applies warning labels to confirmed false information. However, the ever-evolving nature of Fake News presents a challenge, making it difficult to ensure comprehensive content verification. To address this, Facebook encourages users to report suspected false information through their clear reporting method[7].

Meanwhile, Twitter, although relatively new to the policy of addressing misinformation, has made promising strides in combatting false content. They employ a combination of human review, technology, and collaboration with third-party experts to identify and address misleading information. Twitter's unique feature prompts users to reconsider retweeting articles without reading them, reminding them of the potential for misleading headlines. However, the specifics of the technology and expert partnerships remain unclear, and there is a need for more transparency regarding their approach. The reporting system on Twitter offers assorted options, though it can be confusing when it comes to reporting Fake News or false content.

---

4 https://www.openthebox.io/

5 https://www.getbadnews.com/books/english/

6 https://artifact.news/

7 https://transparency.fb.com/it-it/policies/community-standards/misinformation.

Amidst these challenges, Artifact, a text-based news app driven by AI, presents itself as a possible solution to Filter Bubbles and Clickbaits[8]. Developed by Kevin Systrom, the co-founder of Instagram, Artifact aims to counter the infodemic witnessed during the COVID-19 pandemic. The app asks users to select their topics of interest and official news sources during onboarding, allowing it to tailor the content delivered to individual users. Unlike Facebook, Artifact curates its news sources based on integrity rates analyzed by third-party fact-checking services. Additionally, the app leverages ChatGPT technology to provide easy article summaries for users with time constraints. Nicknamed the "TikTok of news,", because it has the same scrolling feed, but has articles instead of videos, Artifact uses The Transformer machine learning system to offer this scrolling feed of news in text format. This system, initially developed by Google in 2017, considers factors beyond clicks, such as dwell/read time and shares, to avoid serving clickbait and filter bubbles. The app's reinforcement learning algorithm, Epsilon-Greedy, ensures users are exposed to a diverse range of recommendations. According to the founders, Artifact also provides a reporting system to combat clickbait and misleading headlines. However, being new in the social media landscape, little information is available, but it would be useful to continue analyzing and studying it, to see if and how it would differentiate itself from other similar platforms.

4.4 Media Policies, Reporting Methods, and Users: Possible Solutions

In this dynamic landscape of social media and news platforms, Facebook, Twitter, and Artifact are continuously evolving their policies and technologies to tackle misinformation while upholding freedom of speech. Their approaches may differ, but all three platforms are committed to addressing the challenges posed by Fake News and ensuring users have access to reliable information. Media literacy education, and information security awareness should be improved, for users of all ages and education levels [20]. Social media could play a key role. From one hand, if all social media adopted small features like a clear reporting system, the prompt feature on Twitter, or warning/labels on those contents that need to be verified yet, users could be slowly and automatically educated to the use of media. On the other hand, a simple cooperation with social media may not be enough since they may be moved from economic interests, so empowering users could work [20]. Additionally, establishing a national-specific social media platform like Artifact, hosting verified and trusted sources, could provide an official, non-political online channel for news consumption and sharing, akin to traditional newspapers and newscasts. Furthermore, considering the registration requirement for journalists, a similar approach could be implemented for online news platforms, ensuring transparency, and eliminating anonymity. This, in conjunction with algorithmic and AI regulations, forms a comprehensive strategy to address misinformation from a regulatory standpoint. However, it is crucial to complement regulatory efforts with technological advancements, such as the development of automated tools for Fake News detection.

[8] https://techcrunch.com/

5 Conclusion and Future Work

Facing the regulatory challenges of Fake News is complex, as evidenced by the analysis of literature and comparative methodologies. Terminological issues, including the definition of terms and the emergence of digital neologisms, pose translation challenges and hinder compliance and harmonization across member states and international borders. This lack of regulation is compounded by the risk of noncompliance with the right to freedom of speech and expression. Furthermore, the European attempt to cooperate with social media platforms has revealed that compliance with agreements depends on economic interests. To address this, enhanced social media regulations are needed, particularly regarding algorithms and the handling of misleading content. By doing so, social media platforms can contribute to the improvement of digital media literacy and educate users effectively. Our future work will focus on the multilingual version of the ontology presented in this paper. We will analyze how compliance among regulations and translation processes can be enhanced. We will also focus on the way the technologies of ChatGPT can support us in translation, Fake News management, and in the analysis of new news app like Artifact.

References

1. Viola, C., Toma, P., Manta, F., Benvenuto, M.: The more you know, the better you act? Institutional communication in Covid-19 crisis management. Technol Forecast Soc Change. **170**, 120929 (2021). https://doi.org/10.1016/j.techfore.2021.120929
2. Berghel, H.: Oh, what a tangled web: russian hacking, fake news, and the 2016 US presidential election. Computer (Long Beach Calif) **50**, 87–91 (2017). https://doi.org/10.1109/MC.2017.3571054
3. Carrieri, V., Madio, L., Principe, F.: Vaccine hesitancy and (fake) news: quasi-experimental evidence from Italy. Health Econ. (United Kingdom). **28**, 1377–1382 (2019). https://doi.org/10.1002/hec.3937
4. Belova, G., Georgieva, G.: Fake News as a Threat to National Security. In: International conference Knowledge-based Organization, vol. 24, pp. 19–22 (2018).https://doi.org/10.1515/kbo-2018-0002
5. Richter, A.: Fake news and freedom of the media. J. Int. Media Entertainmet Law **8**, 1–34 (2020)
6. Goldberg, D.: Responding to "fake news": is there an alternative to law and regulation? Southwest Law Review. **47**, 417–447 (2018)
7. Xu, F., Sheng, V.S., Wang, M.: A unified perspective for disinformation detection and truth discovery in social sensing: a survey. ACM Comput. Surv. **55**, 1–33 (2021). https://doi.org/10.1145/3477138
8. Durach, F., Baragaoanu, A., Nastasiu, C.: Tackling dsinformation: EU regulation of the digital space. Rom. J. Eur. Aff. **20** (2020)
9. Falduti, M., Griffo, C.: Modeling Cybercrime with UFO: an Ontological Analysis of Non-Consensual Pornography Cases. In: Ralyté, J., Chakravarthy, S., Mohania, M., Jeusfeld, M.A., Karlapalem, K. (eds) Conceptual Modeling. ER 2022. LNCS, vol. 13607, pp. 380–394. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-17995-2_27
10. Smal, L., Lösch, A., Van Genabith, J., Giagkou, M., Declerck, T., Busemann, S.: Language Data Sharing in European Public Services-Overcoming Obstacles and Creating Sustainable Data Sharing Infrastructures (2020)

11. Castano, S., Falduti, M., Ferrara, A., Montanelli, S.: The LATO knowledge model for automated knowledge extraction and enrichment from court decisions corpora. Inf. Syst. **106**, 101842 (2022)
12. Castano, S., Falduti, M., Ferrara, A., Montanelli, S.: The CRIKE data-science process for legal knowledge extraction discussion paper. In: Italian Symposium on Advanced Database Systems: 16th–19th June (2019)
13. Castano, S., Falduti, M., Ferrara, A., Montanelli, S.: Law data science and ethics: the CRIKE approach. In: Proceedings of the 1st International Workshop on Processing Information Ethically co-located with 31st International Conference on Advanced Information Systems Engineering, pp. 1–9 (2019)
14. EPRS-ComparativeLaw, europarleuropaeu: La libertà di espressione, una prospettiva di diritto comparato Unione europea. (2019)
15. Longo, L., Goebel, R., Lecue, F., Kieseberg, P., Holzinger, A.: Explainable artificial intelligence: concepts, applications, research challenges and visions. In: Holzinger, A., Kieseberg, P., Tjoa, A.M., Weippl, E. (eds.) CD-MAKE 2020. LNCS, vol. 12279, pp. 1–16. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-57321-8_1
16. Amelio, A., Bonifazi, G., Janusz, A., Briguglio, L., Morpurgo, F., Occhipinti, C.: A multilayer network-based approach for interpreting and compressing convolutional neural networks. A governance and assessment model for ethical artificial intelligence in healthcare. Cognit. Comput. **15**, 1–29 (2022)
17. Eberle, E.J.: Issue 1 Symposium: Methodological Approaches to Comparative Law Article 2 Winter (2011)
18. Zecca, D.: Tutela dell'integrità dell'informazione e della comunicazione in rete: obblighi per le piattaforme digitali fra fonti comunitarie e disciplina degli Stati membri. DPCE Online. 37 (2018)
19. Birritteri Emanuele: Punire la disinformazione. Diritto Penale Contemporaneo - Rivista Trimestrale. 304–334 (2021)
20. Vese, D.: Governing fake news: the regulation of social media and the right to freedom of expression in the era of emergency. Eur. J. Risk Regul. **13**, 477–513 (2022). https://doi.org/10.1017/err.2021.48
21. Kellner, D.: Cultural Studies, Multiculturalism, and Media Culture. SAGE Publications, Thousand Oaks
22. Chaudhary, S., Gkioulos, V., Katsikas, S.: Developing metrics to assess the effectiveness of cybersecurity awareness program, (2022). https://doi.org/10.1093/cybsec/tyac006
23. Bitton, R., Boymgold, K., Puzis, R., Shabtai, A.: Evaluating the information security awareness of smartphone users (2019)