# Community resilience in Metaverse through adopting digital disinformation detection strategies

**Alin ZAMFIROIU, Ella-Magdalena CIUPERCĂ, Simona VOICU,**

**Carmen-Elena CÎRNU, Adrian-Victor VEVERA**

National Institute for Research and Development in Informatics - ICI Bucharest, Romania

alin.zamfiroiu@ici.ro, ella.ciuperca@ici.ro, simona.voicu@ici.ro,

carmen.cirnu@ici.ro, victor.vevera@ici.ro

**Abstract:** An impressive number of researchers have studied the resilience of communities, which is not surprising considering that the community represents an entire world that can be approached through different angles and lenses by different scientific areas. Soon digital space was perceived as a new playground for deception, and Metaverse may be the next target due to the increased exposure of users and their information to vulnerabilities and risks. This article explores the key strategies presented in the literature to combat online disinformation within the Metaverse, indicating the most effective suggestions to enhance community resilience. Emphasis is placed on the importance of equipping community members with tools and strategies to ensure physical well-being and quality of life. Additionally, individuals should learn and apply concrete strategies to maintain comfort and safety in virtual environments, which can be metaphorically referred to as "techniques of digital self-defence in the Metaverse". To facilitate the education of individuals for this purpose, a prototype application where concrete scenarios can be provided for testing has been proposed.

**Keywords:** Community Resilience, Metaverse, Online Disinformation, Digital Self-defence.

# Reziliența comunității în Metavers prin adoptarea de strategii pentru detectarea dezinformării digitale

**Rezumat:** Un număr impresionant de cercetători s-au concentrat asupra studiului rezilienței comunităților, ceea ce nu este surprinzător având în vedere că orice comunitate reprezintă un întreg univers complex care poate fi abordat din diferite unghiuri și perspective științifice. La rândul său, spațiul digital a fost perceput ca un teren fertil pentru înșelăciune și, de aceea, Metavers-ul ar putea deveni următoarea țintă pentru diferite influențe și păcăleli din cauza expunerii crescute a utilizatorilor și a informațiilor specifice acestora la vulnerabilități și riscuri. Acest articol explorează principalele strategii prezentate în literatură pentru a combate dezinformarea online în cadrul Metavers-ului, identificând cele mai eficiente sugestii pentru a spori reziliența comunității. Se pune accent pe importanța dotării membrilor comunității cu instrumente și strategii care să le asigure bunăstarea și calitatea vieții. Oamenii ar trebui să învețe să aplice strategii concrete pentru a menține confortul și siguranța în mediile virtuale, care pot fi denumite metaforic "tehnici de autoapărare digitală în Metavers". Pentru a facilita educarea indivizilor în spiritul acestora, a fost propusă o aplicație prototip în care pot fi oferite scenarii concrete pentru testare.

**Cuvinte cheie:** Metavers, dezinformare online, autoapărare digitală, reziliență, comunitate.

## 1. Introduction

Virtual communities have found their greatest success in games like Second Life (launched in 2003), where users can create avatars to interact with each other in a virtual world (Jaipong et al., 2023). Although Second Life was designed as an entertainment platform, it was also used to build community resilience by training social skills, education and personal development (Nguyen & Nof, 2018). Community members can explore virtual worlds, strengthen social skills, and participate in support communities for a variety of social issues, such as support groups for people going through difficult times or groups formed around shared interests such as art, business, or education.

Starting with Second Life, online communities constantly emerged being primarily focused on discussion forums, support groups, or platforms for people with shared economic or professional interests. Nowadays cyber-environment is different from the virtual community that Howard Rheingold coined in 1993 to describe groups of people who gather in digital spaces to

https://doi.org/10.33436/v34i3y202402

communicate, collaborate and interact. In addition to their motivations for joining these communities, members are also expected to follow a set of norms and rules that help to create a positive and functional online environment, including online etiquette, coexistence guidelines, and conflict resolution mechanisms.

Metaverse is the culmination of digitization and technological advancement, incorporating disruptive cutting-edge technologies such as artificial intelligence (AI), blockchain, virtual reality (VR), and augmented reality (AR). It is the next level of social interaction, including social media and taking it to the next level. Digital twins combined with virtual reality will allow near-real experiences, avoiding many disadvantages such as travel time loss, accidents during simulations etc., while maintaining the benefits of experimentation (Siyaev & Jo, 2021). It is believed that Metaverse will become a place where entertainment, business and social life converge, generating a market of over $1.5 trillion by 2030 (PwC report, 2019). The Metaverse surged into prominence when the CEO of the largest social media platform proclaimed it as the next revolutionary development following the Internet. While a universally accepted definition is yet to be established, the Metaverse is broadly understood to integrate IoT, AR, VR, XR and 3D technologies. It is also referred to as Web 3.0.

Many studies (Perkowitz et al., 2003; Visconti, 2022), have investigated the complex interactions between real-world communities and the online world, exploring how the digital space influences the behaviour of its members and how these behaviours can be transformed by the digital realm, leaving a unique mark on community life.

Individuals targeted in an online environment can lose not only personal, but also critical information and even parts of their own ideas and identity. Which means that in fact they may lose their cognitive autonomy because of digital disinformation. These forms of influence have significant consequences, including shaping public opinion, elections and causing harm to individuals or communities. They undermine trust in institutions, polarize public discourse and contribute to the erosion of democratic processes. Moreover, they can lead to tangible real-world consequences, such as influencing public health behaviours, inciting violence, or manipulating financial markets.

In Metaverse, social presence is transposed into a virtual dimension, where users can interact through personalized avatars (Kim et al., 2023). These digital representations allow individuals to express unique identities and participate in common activities, from business meetings to virtual concerts. Communication becomes more expressive, going beyond the limits of text or traditional video calls, offering a wide range of facial expressions, gestures, and virtual body language.

Undoubtedly, people's online behaviour can be significantly different from their real-world behaviour, as described by Proteus effect. According to Proteus effect, people's behaviour in virtual worlds is influenced by the characteristics of their avatars, which they choose, understanding that other individuals in the digital space are evaluating certain behavioural traits in specific ways (Yee & Bailenson, 2007). In order to enhance their standing among others in the Metaverse, individuals leverage common trends in assessing others and fine-tune their avatars to obtain maximum benefits in the digital realm.

In these conditions, one major challenge of virtual interactions is the simplified assessment of individuals, as many cues, such as nonverbal and paraverbal language, emotions, past relational experiences, and contextual positioning, are absent. As a result, people are evaluated based on limited information, and the assessment they receive based on the few pieces of information is usually wrong or meaningless. But people are also aware of this, therefore they provide precisely those unverifiable pieces of information that allow them to associate with the most favourable identities. This way, they can gain online what they often struggle to achieve in the real world.

The question of how much truth and authenticity is encouraged online arise. What are the chances for an authentic person to make a good impression in these virtual communities? How is it possible to identify deception in online communities? How can a specific virtual community protect itself from the intrusion of fake individuals and maintain its status quo? How can the resilience of communities be enhanced, especially those that face initial challenges in forming their

identity? And how to translate all these answers in Metaverse? In this article, the main goal is to examine individual and community online resilience from the perspective of digital deception and to suggest a few approaches aimed at enhancing overall well-being.

## 2. Related work

As the recent pandemics shows, the online environment facilitates remote collaboration by removing physical and temporal barriers. Teams can work together in virtual workspaces, manipulating 3D objects and simulating real-world scenarios. This ability to collaborate in a controlled and flexible environment can accelerate innovation and improve the efficiency of work processes (Brand et al., 2023). It is said that Metaverse will be a disrupting technology that will unleash its potential in many fields. Education is one of the most dynamic social sectors, assimilating and applying innovation. Therefore, it is expected that the educational sector may benefit enormously from Metaverse. Students may participate in interactive lessons in virtual environments that simulate historical situations, scientific experiments, or cultural excursions. By immersing themselves in the educational content of Metaverse, learning becomes an engaging and memorable experience, increasing information retention and student engagement.

While digital literacy and education at large is considered to be a predictor of identifying deception in offline and online environment (Gaillard et al., 2021; Pandey, 2018; Sarno & Black, 2024), there are many recommendations that education for using Metaverse should include a series of items that make people more susceptible to identify disinformation and fight against it. Detecting and combating digital fraud in the online environment has been and is addressed in several scientific works. Among the most important authors who have dealt with the analysis of methods and strategies to combat online fraud are: Claire Wardle (Wardle & AbdAllah, 2023), Hany Farid (Booth et al., 2024), Sander van der Linden (Ecker et al., 2024) or Filippo Menczer (Jahn et al., 2023), who proposed different methods and strategies for detecting and combating digital fraud in the online environment, Figure 1.

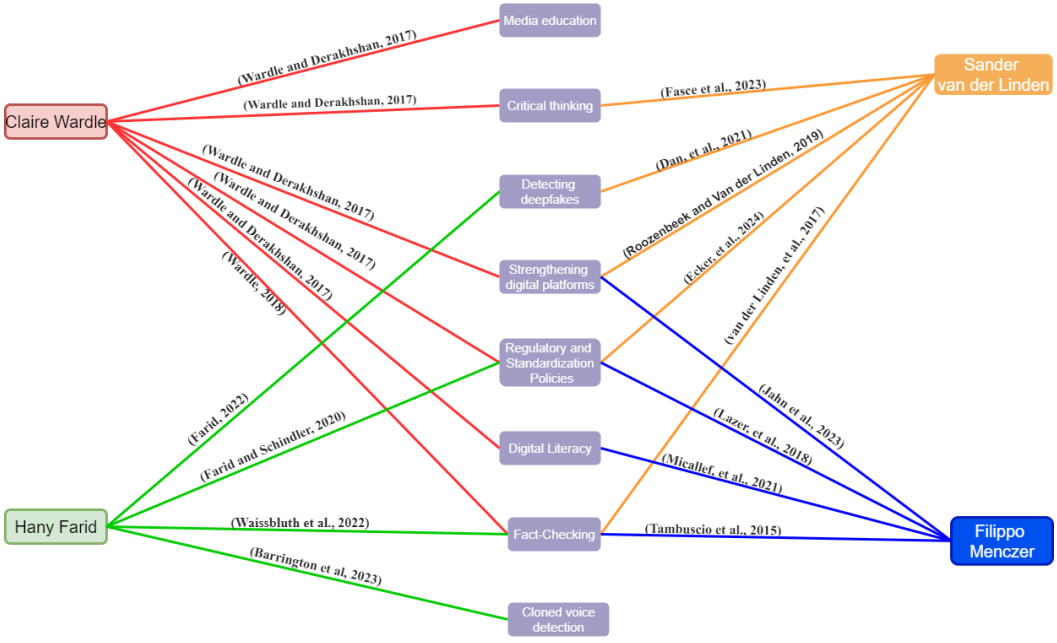


**Figure 1.** Strategies presented in the literature used for detecting and combating digital deception

For the Media Education strategy, according to (Wardle & Derakhshan, 2017), media organizations should collaborate and produce policies and news literacy segments, improving the quality of content.

According to (Fasce et al., 2023) the human person creates two hypotheses that will be believed or not. These hypotheses are created based on group belongingness, perceived threat on

that field and or intergroup anxiety. In this way a person should have both hypotheses in the same direction to believe a supposition. If the second hypothesis is different than the critical thinking will remove that supposition.

For the Strengthening Digital Platforms strategy, according to Wardle & Derakhshan (2017) technology organizations should collaborate to provide transparent criteria for any algorithmic changes that down-rank content, to eliminate financial incentives, to pay attention to audio/visual forms of mis- and dis-information and provide metadata to trusted partners and/or verification tools.

In (Barrington et al., 2023) three approaches for classifying speech as synthetic or real, and for identifying the underlying synthesis architecture are described. Three techniques for differentiating a real from a cloned voice designed to impersonate a specific person are described in this paper.

Detecting digital deception is an ongoing process that involves critical thinking and a healthy level of scepticism. By applying these strategies and remaining vigilant, users can become more adept at identifying deceptive content and protecting themselves from potential harm.

## 3. Detecting digital deception in online environments

Digital deception in online environments refers to a broad range of deceptive or fraudulent activities that occur over the internet (Mavlanova, 2008). It involves various tactics used by malicious actors to trick, manipulate, or defraud individuals or organizations in the digital realm.

In the context of the online environment, the speed and scale at which information spreads make it easier for digital deception, misinformation and disinformation to be disseminated and amplified. Identifying and combating these issues require critical thinking, media literacy, fact-checking, and the use of technology to detect and counter deceptive content. Additionally, social media platforms, news organizations, and governments play a role in curbing the spread of false information and digital deception.

Despite the various concepts reflecting different dimensions of deception in social sciences, the perspective proposed by Hameleers et al. (2022) concerning misinformation and disinformation in digital communication has been adopted. These forms of information can have significant consequences, including shaping public opinion, influencing elections, and causing harm to individuals or communities. These concepts encompass most of the aspects discussed below, with the primary distinguishing criterion being the intent of the information source:

- **Misinformation** refers to inaccurate or false information that is spread without the intention to deceive. It can arise from various sources, including misunderstandings, rumours, or genuine mistakes. In the context of digital communication, misinformation can quickly spread through social media, websites, blogs and other online platforms. Examples of misinformation include false news stories, fabricated quotes and misleading statistics;
- **Disinformation**, on the other hand, is intentionally false or misleading information that is spread with the purpose of deceiving or manipulating others. Disinformation campaigns are often coordinated efforts to spread false narratives, sow confusion, or advance specific agendas. They can be carried out by individuals, organizations, or even state-sponsored actors. Disinformation can take many forms, such as fabricated news articles, fake images or videos, and manipulated facts or statistics.

Both misinformation and disinformation can have significant social, political and economic implications. They can undermine trust in institutions, polarize public discourse, and contribute to eroding democratic processes. They can also have real-world consequences, such as influencing public health behaviours, inciting violence, or manipulating financial markets by propaganda, rumour, or other techniques (Guess & Lyons, 2020). Online disinformation is often used to manipulate public opinion and influence people's perceptions of specific topics or events. By spreading false or distorted information, malicious actors can create or reinforce certain beliefs or

fuel negative sentiments toward particular groups or ideologies. Also contributing to societal division and polarization by promoting extreme, controversial, or inflammatory content, it amplifies existing social and political tensions, creating communication barriers between different groups and fuelling internal conflicts. During elections or referendums, spreading false information can undermine trust in the integrity of these processes and impact democratic outcomes. In Figure 2, digital deception techniques, as presented in the literature, were synthesized.

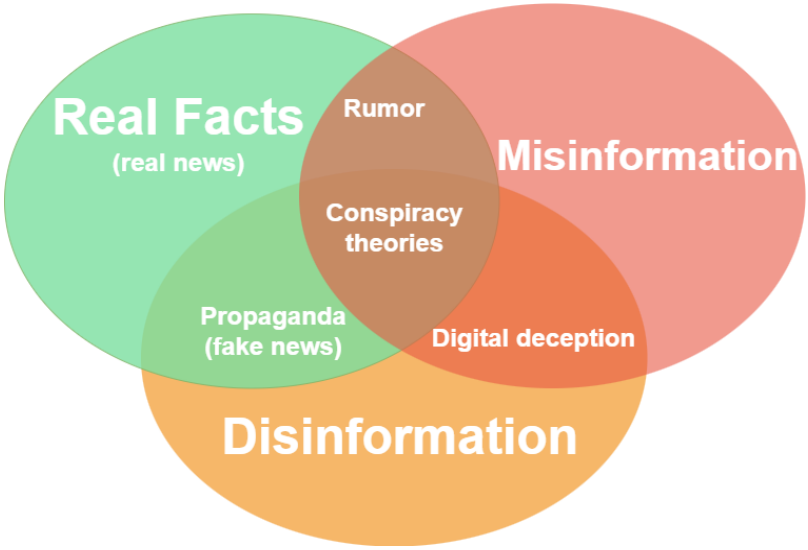


**Figure 2.** Digital deception techniques in online environments

The World Health Organization (WHO) and its partners acknowledge that online misinformation has the potential to spread further, faster, and sometimes deeper than the truth. On some social media platforms, false news is 70% more likely to be shared than accurate news. To counteract this phenomenon, the WHO has taken a series of actions in collaboration with technology companies to promote accurate information and combat misinformation (WHO, 2024).

Misinformation can manipulate voters and influence political decisions in ways that may not reflect the true will of the population. Sometimes, online disinformation is used purely for financial gain or personal advantage. By creating viral content or promoting fraudulent products or services, disinformers can earn money or gain power and influence in the online space. Particularly in the field of health, online disinformation can have serious consequences. Spreading false information about medical treatments or vaccines can jeopardize people's health and safety, contributing to increased concerns and reducing vaccine coverage (Kreps, 2020).

Emotional language can indeed be a powerful tool for spreading misinformation. When information triggers strong emotions, whether it is fear, anger, or happiness, people are often more inclined to share it without critically evaluating its accuracy. It is essential to approach information with calmness and vigilance and analyse it with a rational mindset, especially when emotions are deliberately manipulated to influence opinions or beliefs.

## 4. Implications of Digital deception in Metaverse and community resilience

The development of the Metaverse will transform online communities in innovative ways, thanks to technological advances such as avatar persistence, uniqueness, and interconnectivity. To fulfil its social mission, the digital world, which will soon include the Metaverse, the newest platform of the present and future (Nath, 2022), must reinvent trust between people, as well as trust in hardware and infrastructure (Barbu, 2016). To enhance trust, the Metaverse should be perceived as a faithful representation of the physical world. New rules should prioritize the latest technologies instead of old traditional regulations, as they may dynamically alter the digital landscape and make even the newest regulations obsolete in a short period of time.

Interactivity specific to the Metaverse, its realism, and direct interactions between users create a conducive environment for the emergence of much more severe effects of disinformation and misinformation.

In this environment, where real-world rules are intertwined, the real and digital worlds are expected to become increasingly integrated, with individuals working and living simultaneously in both realms. Consequently, it is anticipated that social behaviour will be governed by new rules blending the regulations of both worlds. Ensuring behavioural consistency may become one of the most significant challenges for regulators. Otherwise, people may experience cognitive dissonance and they will be subjected to different mental disorders.

The safety of social relationships in Metaverse can be compromised from several angles. Anonymity can be a major source of digital deception, providing a perfect shield for people to engage in manipulation and disinformation, from spreading misinformation to orchestrating large-scale disinformation and propaganda campaigns.

The strategies discussed earlier, useful for combating misinformation in the online environment in general, will be applied on this new platform as well. However, a discussion regarding how they should be optimized is necessary to address the unique characteristics of the Metaverse, previously unseen in other types of digital environments, to foster real and functional online free-of-deception communities.

Given the significant role of artificial intelligence in building the Metaverse, it would be ideal for users to *check facts* in real-time using dedicated AI algorithms. Additionally, assigning virtual trust badges to verified information would facilitate easier navigation through the vast expanse of data and information.

User *digital education* can be considered a solution to many of humanity's current and future problems. In the Metaverse, it is easy to estimate that a good understanding of how the new platform works, the typology of users, the interests they may represent, as well as a personal toolkit for combating disinformation through personal life experiences and accumulated knowledge, will help reduce misinformation. Digital literacy can be an end in itself or integrated into various gaming or social experiences on the platform. The development of *critical thinking* is closely linked to digital education, as it is important for users to better evaluate information and identify potential fake news. *Media education* for Metaverse should include educational modules for users training to recognize visual and audio manipulations of VR and AR environments.

The specifics of the Metaverse, compared to previous versions of digital environments, draw attention to the accelerated obsolescence of rules, confronting us with a new psychosocial dynamic. In this context, new rules emerge while people have not yet become accustomed to the previous ones. Without being internalized, the old rules become obsolete, and new sets of rules take their place. The rapid pace of this process can result in the emergence of digital anarchy, defined as the absence of an appropriate regulatory framework - not necessarily due to a lack of regulation, but because of the need for constant adaptation and continuous change in a dynamic environment. It is important to notice that responsibility is not a voluntary matter. This is why Metaverse platform should be properly and sufficiently *regulated* through international collaboration to impose sanctions on those who spread disinformation and misinformation. An important part of data protection and confidentiality is assured through *strengthening digital platforms* using communication encryption and implementation of anomaly detection algorithms which may help prevent the spreading of disinformation and misinformation.

Although avatars will be used in the Metaverse, making interaction with one's own image and voice optional, there is the possibility of using one's own voice. Therefore, when users interact directly, it is important to have technologies such as advanced voice recognition algorithms to prevent cloned *voice detection*. Also, artificial intelligence algorithms can analyse Metadata and video pixels to detect deep fakes and classify them based on their intended negative or positive purposes (Stanciu & Ciuperca, 2024). These checks should ideally be performed before publication, and highly reliable detections should be certified with specific badges.

What would be the prototype of the ideal community of Metaverse? While there is no concrete research on ideal communities in the Metaverse yet, based on the literature, it can be hypothesized that such a community would have the following characteristics:

- Members with a strong sense of belonging, which could be measured using specific tools;
- The ability to adapt to changes in the real and virtual worlds, and to adjust its goals and rules as needed;
- A clear framework for providing practical and emotional support to members who are facing problems in the Metaverse;
- Clear and well-established communication channels that allow members to easily share information with each other;
- Clear mechanisms for making decisions and resolving conflicts.

According to the characteristics of virtual communities in the Metaverse, the following sources of resilience can be identified:

1. Establishing clear rules to govern people's interactions on the platform;
2. Regulating all types of interactions, including the use of artificial intelligence to interpret data collected on the platform and warn of potential problems;
3. Securing sensitive data in encrypted spaces using highly efficient technology;
4. Using intrusion monitoring and detection systems;
5. Planning ahead for cyberattacks to minimize damage and facilitate recovery.

As Metaverse is perceived as a new playground for disinformation and deception, each member of the community should be aware of the previous strategy to guarantee physical well-being and quality of life - they are metaphorically named "*techniques of digital self-defence in Metaverse*".

## 4. Use case: ICI EDscape Room

To assess the extent to which community resilience can be enhanced by practicing the ability to identify deception strategies, a virtual escape room was utilized, originally designed to test ICDL (International Computer Driving License) knowledge. This app is developed as a prototype based on recommendations regarding the imagined form of Metaverse and was designed to improve users' ICT knowledge. For the current issue, ICI EDscape Room with a knowledge quiz regarding the most important methods of identifying misinformation online was used. The users had to identify the correct answer for a set of 30 questions, with different degrees of difficulty, each correct answer bringing 3 points out of 100 possible. As in the traditional educational scoring system, users receive 10 points from the beginning and if they manage to achieve a total of 70 points, they get the opportunity to access the next room. In the next room, they are exposed to different scenarios containing correct and incorrect information based on the previous methods and they must evaluate information credibility.

This application aims to implement some of the strategies presented in the literature (Figure 1). The following strategies are considered to be applied:

- Education – the purpose of the game is educational, with the intent of using this serious game to help users accurately understand how certain systems function, avoiding incorrect assumptions;
- Critical Thinking – in the quiz section, the questions are built in that manner that the user have to think critically for the answers and after this section in the rooms dedicated to construct the computer, phone or other systems, they have to apply the knowledge form the quiz questions; In this way we apply the results of (Fasce et al.,

2023) that the user will have the first hypothesis in the quiz, and second hypothesis in the practical way;

- Fact checking – because the user will check if the systems are correctly constructed. If the systems are wrong the application will advise the user how the system can be improved.

## 5. Conclusions

While the emergence of the Metaverse represents a profound shift in digital interaction, building trust during navigating this new platform is a paramount concern. The most important strategies for detecting and combating digital deception in online environments have been analysed. These strategies are also applicable to the Metaverse. Additionally, a prototype application has been developed to test some of these strategies in combating digital deception for users.

Looking ahead, the prototype of an ideal Metaverse community would include inclusivity, adaptability and resilience. Such a community and the applications developed in this environment would not only prioritize the well-being and safety of its members but also facilitate transparent communication, robust conflict resolution mechanisms, and continuous adaptation to evolving technological landscapes in the field. By striving towards these ideals and leveraging innovative solutions, the full potential of the Metaverse will be unlocked as a transformative force for social connectivity and progress.

### Acknowledgement

The work presented in this paper is supported by the Core Program within the National Research Development and Innovation Plan 2022-2027, carried out with the support of MCID, project no. 23380601, “Advanced research in the Metaverse and emerging technologies for the digital transformation of society”.

## REFERENCES

Barbu, D. C. (2016) Îmbunătăţirea protecţiei infrastructurilor critice din sectorul TIC prin creşterea rezilienţei (Improving the protection of critical IT&C sector infrastructures by increasing resilience). *Revista Română de Informatică şi Automatică*. 26(4), 29-34.

Barrington, S., Barua, R., Koorma, G. & Farid, H. (2023) Single and multi-speaker cloned voice detection: From perceptual to learned features. In *2023 IEEE International Workshop on Information Forensics and Security (WIFS), 4-7 December, 2023, Nürnberg, Germany*. IEEE. pp. 1-6.

Booth, E., Lee, J., Rizoiu, M. A. & Farid, H. (2024). Conspiracy, misinformation, radicalisation: understanding the online pathway to indoctrination and opportunities for intervention. *Journal of Sociology*. doi:10.1177/14407833241231756.

Brand, D., Knopf, T. & Stumpp, S. (2023) Metaverse and Social Virtual Reality for Online Collaboration. In Moreira, F. & Jayantilal, S. (eds.) *Proceedings of the 18th European Conference on Innovation and Entrepreneurship, ECIE 2023, Part 1, 21-22 September, 2023, Universidade Portucalense, Porto, Portugalia.* pp.131-139.

Dan, V., Paris, B., Donovan, J., Hameleers, M., Roozenbeek, J., van der Linden, S. & von Sikorski, C. (2021) Visual mis-and disinformation, social media, and democracy. *Journalism & Mass Communication Quarterly*. 98(3), 641-664.

Ecker, U. K., Tay, L. Q., Roozenbeek, J., van der Linden, S., Cook, J., Oreskes, N. & Lewandowsky, S. (2024) *Why misinformation must not be ignored*. https://osf.io/8a6cj/download.

Falchuk, B., Loeb, S. & Neff, R. (2018) The social metaverse: Battle for privacy. *IEEE Technology and Society Magazine*. 37(2), 52-61.

Farid, H. & Schindler, H. J. (2020) Deep Fakes. *On the Threat of Deep Fakes to Democracy and Society. Berlin: Konrad Adenauer Stiftung*. Online: https://tohear.info/pdf/CEP-KAS_Deep%20Fakes_062920.pdf.

Farid, H. (2022) Creating, Using, Misusing, and Detecting Deep Fakes. *Journal of Online Trust and Safety*. 1(4), 1-33. doi:10.54501/jots.v1i4.56.

Fasce, A., Adrián-Ventura, J., Lewandowsky, S. & van der Linden, S. (2023) Science through a tribal lens: A group-based account of polarization over scientific facts. *Group Processes & Intergroup Relations*. 26(1), 3-23. doi:10.1177/13684302211050323.

Gaillard, S., Oláh, Z. A., Venmans, S. & Burke, M. (2021) Countering the cognitive, linguistic, and psychological underpinnings behind susceptibility to fake news: A review of current literature with special focus on the role of age and digital literacy. *Frontiers in Communication*. 6, 661801. 1-14. doi:10.3389/fcomm.2021.661801.

Guess, A. M. & Lyons, B. A. (2020) Misinformation, disinformation, and online propaganda. In: Persily, N. & Tucker, J.A. (eds.) *Social Media and Democracy. The State of the Field, Prospects for Reform*. Cambridge University Press. pp.10-33.

Hameleers, M., Brosius, A. & de Vreese, C. H. (2022) Whom to trust? Media exposure patterns of citizens with perceptions of misinformation and disinformation related to the news media. *European Journal of Communication*. 37(3), 237-268. doi:10.1177/02673231211072667.

Jahn, L., Rendsvig, R. K., Flammini, A., Menczer, F. & Hendricks, V. F. (2023) Friction Interventions to Curb the Spread of Misinformation on Social Media. To be published in *Social and Information Networks*. [Preprint] [Accessed: August 2024]. https://doi.org/10.48550/arXiv.2307.11498.

Jaipong, P., Siripipattanakul, S., Sriboonruang, P., Sitthipon, T., Jaipong, P., Siripipattanakul, S., ... & Sitthipon, T. (2023) A review of metaverse and cybersecurity in the digital era. *International Journal of Computing Sciences Research*. 7, 1125-1132.

Kim, D. Y., Lee, H. K. & Chung, K. (2023) Avatar-Mediated Experience in the Metaverse: The Impact of Avatar Realism on User-Avatar Relationship. *Journal of Retailing and Consumer Services*. 73(8), 103382. doi:10.1016/j.jretconser.2023.103382.

Kreps, S. (2020) The role of technology in online misinformation. *Foreign Policy*. 2-7, https://www.brookings.edu/wp-content/uploads/2020/06/The-role-of-technology-in-online-misinformation.pdf [Accessed: August 2024]

Mavlanova, T., Benbunan-Fich, R. & Kumar, N. (2008) Deception tactics and counterfeit deception in online environments. *ICIS 2008 Proceedings*. 105, 1-11. https://aisel.aisnet.org/icis2008.

Nath, S. (2022) Accounting in the Virtual World: Legal & Ethical Challenges in the Metaverse Economy.

Nguyen, W. P. & Nof, S. Y. (2018) Resilience informatics for cyber-augmented manufacturing networks (CMN): Centrality, flow and disruption. *Studies in Informatics and Control*. 27(4), 377-384.

Pandey, N. (2018) Fake news: A manufactured deception, distortion and disinformation is the new challenge to digital literacy. *Journal of Content, Community and Communication*. 4(8), 15-21. doi:10.31620/JCCC.12.18/04.

Perkowitz, M., Philipose, M. & McCarthy, J. F. (2003) Utilizing online communities to facilitate physical world interactions. In *The International Conference on Communities and Technologies*. pp. 1-6.

PwC Report (2019) *PwC Annual Report* https://www.pwc.com/my/en/publications/2019/pwc-annual-report-2019.html

Roozenbeek, J. & Van der Linden, S. (2019) Fake news game confers psychological resistance against online misinformation. *Palgrave Communications*. 5(1), 1-10.

Sarno, D. M. & Black, J. (2024) Who Gets Caught in the Web of Lies?: Understanding Susceptibility to Phishing Emails, Fake News Headlines, and Scam Text Messages. *Human Factors*. 66(6), 1742-1753. doi:10.1177/00187208231173263.

Stanciu, A. & Ciupercă, E. (2024) Can Deepfakes Benefit from Metaverse in an Era of Disinformation? Insights from a Systematic Review. *IFAC-PapersOnLine*. 58(3), 61-65. doi:10.1016/j.ifacol.2024.07.125.

Tambuscio, M., Ruffo, G., Flammini, A. & Menczer, F. (2015) Fact-checking effect on viral hoaxes: A model of misinformation spread in social networks. In *Proceedings of the 24th international conference on World Wide Web*. pp. 977-982.

Van der Linden, S., Maibach, E., Cook, J., Leiserowitz, A. & Lewandowsky, S. (2017) Inoculating against misinformation. *Science*. 358(6367), 1141-1142.

Visconti, R. M. (2022) From physical reality to the Metaverse: a Multilayer Network Valuation. *Journal of Metaverse*. 2(1), 16-22.

Waissbluth, E., Farid, H., Sehgal, V., Peshin, A. & Afroz, S. (2022) Domain-Level Detection and Disruption of Disinformation. *arXiv preprint arXiv:2205.03338*.

Wardle, C. & AbdAllah, A. (2023) The Information Environment and Its Influence on Misinformation Effects. In: Purnat, T.D., Nguyen, T. & Briand, S. (eds). *Managing Infodemics in the 21st Century: Addressing New Public Health Challenges in the Information Ecosystem*. Cham: Springer International Publishing, pp.41-51.

Wardle, C. & Derakhshan, H. (2017) *Information disorder: Toward an interdisciplinary framework for research and policymaking*. 27, pp. 1-107. Strasbourg: Council of Europe.

Wardle, C. (2018) Information disorder: The essential glossary. *Harvard, MA: Shorenstein Center on Media, Politics, and Public Policy, Harvard Kennedy School*.

World Health Organization, Combatting misinformation online, https://www.who.int/teams/digital-health-and-innovation/digital-channels/combatting-misinformation-online [Accessed: 25.04.2024]

Yee, N. & Bailenson, J. (2007) The Proteus Effect: The Effect of Transformed Self-Representation on Behavior. *Human Communication Research*. 33(3), 271–290. doi:10.1111/j.1468-2958.2007.00299.x.

**Alin ZAMFIROIU** works as a Senior Researcher at National Institute for Research & Development in Informatics -ICI Bucharest and associate professor at the Department of Economic Informatics and Cybernetics at the Bucharest University of Economic Studies, Bucharest. He has published as an author and co-author of journal articles and presented scientific works at important conferences of the field.

**Alin ZAMFIROIU** lucrează ca cercetător științific gradul I la Institutul Național de Cercetare-Dezvoltare în Informatică - ICI București și este conferențiar universitar la Departamentul de Informatică Economică și Cibernetică al Academiei de Studii Economice București. A publicat în calitate de autor și coautor articole în jurnale, și a prezentat lucrări științifice la conferințe importante în domeniu.

**Ella-Magdalena CIUPERCĂ** is involved in various research projects ranging from critical infrastructure protection to cybersecurity skills training in cyber ranges, open science, Metaverse applications, predominantly contributing through the scientific assurance of social and security dimensions. The recognition that she obtained is proved through her activity as a member of scientific and editorial committees of several scientific journals and events in the field.

**Ella-Magdalena CIUPERCĂ** este implicată în diverse proiecte de cercetare, care variază de la protecția infrastructurilor critice și formarea competențelor în securitate, până la dezvoltarea aplicațiilor în Metavers, contribuția sa principală fiind pe dimensiunea asigurarea expertizei științifice a dimensiunilor sociale și de securitate. Recunoașterea activității sale este evidențiată de participarea în calitate de membru al comisiilor științifice și editoriale ale mai multor reviste și conferințe de prestigiu din domeniu.

**Simona VOICU** has over 10 years of experience in Communication and Social Media, being involved in various projects in the central administration, but also in the private sector. Since 2017 she carries out the activity within the National Institute for Research - Development in Informatics - ICI Bucharest. The main areas of interest for the research activity include: eHealth, business models, strategic communication and digital diplomacy.

**Simona VOICU** are o experiență de peste 10 ani în domeniul comunicării și al social media, fiind implicată în diverse proiecte atât în administrația centrală, cât și în sectorul privat. Din anul 2017 își desfășoară activitatea în cadrul Institutului Național de Cercetare-Dezvoltare în Informatică - ICI București. Principalele sale domenii de interes pentru activitatea de cercetare includ: eHealth, modele de business, comunicare strategică și diplomație digitală.

**Carmen-Elena CÎRNU** is the Scientific Director, Senior Researcher I and member of the Scientific Council of the National Institute for Research and Development in Informatics - ICI Bucharest. She received her Ph.D. in Philosophy in 2011. She is involved in numerous research and development projects in the field of interoperability, cybersecurity and e-Government. She is the author or co-author of numerous articles, studies and research reports.

**Carmen-Elena CÎRNU** este Director Științific, cercetător științific gradul I și membru în Consiliul Științific al Institutului Național de Cercetare-Dezvoltare în Informatică – ICI București. A obținut titlul de doctor în filozofie în anul 2011. Este implicată activ în numeroase proiecte de cercetare-dezvoltare în domeniile interoperabilității, securității cibernetice și e-guvernării. de asemenea este autor și coautor al unui număr semnificativ de articole, studii și rapoarte de cercetare.

**Adrian-Victor VEVERA** is the General Director, First Degree Scientific Researcher and member of the Scientific Council of the National Institute for Research and Development in Informatics. Mr. Vevera holds a Ph.D. in military and information sciences. He has extensive experience in the field of national security. He has published numerous articles and papers on national and international security issues, energy security, cybercrime, critical infrastructure protection, and has been the coordinator of numerous projects of national interest.

**Adrian-Victor VEVERA** este Director General, cercetător științific gradul I și membru în Consiliul Științific al Institutului Național de Cercetare-Dezvoltare în Informatică – ICI București. Este doctor în științe militare și informații. A publicat numeroase articole și lucrări pe teme de securitate națională și internațională, securitate energetică, criminalitate informatică, precum și protecția infrastructurilor critice. A coordonat multiple proiecte de interes național.