frontiers | Frontiers in Computer Science

TYPE Original Research
PUBLISHED 30 October 2024
DOI 10.3389/fcomp.2024.1462250

OPEN ACCESS

EDITED BY
Peter Kieseberg,
St. Pölten University of Applied Sciences, Austria

REVIEWED BY
Elochukwu Ukwandu,
Cardiff Metropolitan University, United Kingdom
Dimiter Velev,
University of National and World Economy, Bulgaria

*CORRESPONDENCE
Simona-Nicoleta Vulpe
✉ simona.vulpe@drd.unibuc.ro

RECEIVED 10 July 2024
ACCEPTED 18 October 2024
PUBLISHED 30 October 2024

CITATION
Vulpe S-N, Rughiniș R, Țurcanu D and Rosner D (2024) AI and cybersecurity: a risk society perspective.
Front. Comput. Sci. 6:1462250.
doi: 10.3389/fcomp.2024.1462250

AI and cybersecurity: a risk society perspective

Simona-Nicoleta Vulpe[1]*, Răzvan Rughiniș[2], Dinu Țurcanu[3] and Daniel Rosner[4]

[1]Faculty of Sociology and Social Work, University of Bucharest, and Research Institute of the University of Bucharest (ICUB), Bucharest, Romania, [2]Faculty of Automatic Control and Computers, National University of Science and Technology POLITEHNICA Bucharest, and The Academy of Romanian Scientists, Bucharest, Romania, [3]Faculty of Electronics and Telecommunications and National Institute of Innovations in Cybersecurity "CYBERCOR", Technical University of Moldova, Chișinău, Moldova, [4]Faculty of Automatic Control and Computers, National University of Science and Technology POLITEHNICA, Bucharest, Romania

Introduction: The rapid evolution of Artificial Intelligence (AI) has introduced transformative potential across various sectors, while simultaneously posing significant cybersecurity risks.

Methods: The aim of this paper is to examine the debates on AI-related cybersecurity risks through the lens of Beck's theory of the risk society. Utilizing thematic content analysis, we explored public discourse on AI and cybersecurity as presented in articles published by WIRED.

Results: Our analysis identified several key themes: the global nature of AI risks, their pervasive influence across multiple sectors, the alteration of public trust, the individualization of risk, and the uneven distribution of AI risks and benefits.

Discussion: The editorial choices in WIRED predominantly favor a functionalist and solutionist perspective on AI cybersecurity risks, often marginalizing the opinions of ordinary individuals and non-Western voices. This editorial bias tends to limit diversity and underrepresent key opposing viewpoints, potentially hindering a more comprehensive and nuanced debate on AI and cybersecurity issues.

KEYWORDS

Artificial Intelligence, cybersecurity, risk society, reflexivity, WIRED, thematic content analysis

1 Introduction

Artificial Intelligence (AI) is a rapidly evolving field that has the potential to drastically alter many facets of society and our daily lives. It includes an extensive range of technologies and applications, all of which help to create intelligent systems that are capable of performing activities that would typically require human intelligence. As AI continues to advance, it will increasingly shape society through a variety of opportunities and challenges.

AI entails a number of risks, ranging from limited AI to artificial general intelligence (AGI). Malicious attacks, malfunctions, and unpredictable behavior in complex systems are common concerns (Page et al., 2018).

Large-scale adoption of AI expands the associated risks, the main concerns being increased structural unemployment and economic inequality (Hutter and Hutter, 2021; Aleshkovski et al., 2022), displacement of human creativity (Bran et al., 2023), potential loss of privacy and growth in cybercrime (Aleshkovski et al., 2022), amplification of hate speech (Rughiniș et al., 2024), and ethical challenges related to AI decision-making (Bergsten and Rivas, 2019). The integration of AI into various sectors may lead to uncontrolled substitution of human roles, posing risks to social cohesion and security (Hutter and Hutter, 2021). Other identified risks include deep automation

bias, insufficient machine learning quality, and the lack of algorithmic accountability (Strauß, 2018). The development of autonomous weapons raises concerns about mass casualties (Chander, 2024). Additionally, the concentration of power and wealth, unknown long-term consequences, and the changing relationship between humans and robots are significant social challenges (Benjamins and Salazar, 2020). Addressing these risks requires a combination of technological solutions, ethical considerations, and policy interventions (Arvind Ashta, 2023; Critch and Russell, 2023).

The way the public engages with technology is transforming social life and shaping perceptions of AI risks. Numerous studies have addressed the topic of AI and how it is perceived by the public in different countries. Budeanu et al. (2023) analyzed public perceptions of the impact AI will have on people's lives in the next 20 years and the impact of this technology in terms of creating new jobs, using Eurobarometer data. According to this analysis, public perception of the future impact of AI on society is determined by cultural specificities and existing national structures. In addition, socio-demographic categories shape people's perceptions of the social impact of AI only to a small extent. Age and gender do not significantly influence public perception of AI. On the other hand, education and social class are significantly associated with the public evaluation of AI, indicating the relevance of social stratification in the context of AI disruptions. People with higher levels of education are more optimistic about the social impact of AI. Similarly, individuals belonging to a higher social class are more likely to have a positive perception of the social impact of AI. In a related study on the German public, the most significant AI risk perceived by respondents, according to Brauner et al. (2023), was cybersecurity issues. However, the authors also pointed out that many people still view AI as a black box, which results in distorted perceptions of related risks.

Previous studies on media content related to AI identified prevailing themes and dominant discourses. News media tend to frame AI through business and technology lenses, with benefits discussed more frequently than risks (Chuan et al., 2019; Sun et al., 2020). However, coverage has become more critical over time, highlighting ethical concerns and data risks like privacy invasion and algorithmic bias (Nguyen, 2023; Nguyen and Hekman, 2024). The framing of AI often involves complex networks of actors, including researchers, businesses, and governments (Zhai et al., 2020). Media portrayal of AI ethics is generally practical but shallow, suggesting a need for more accessible and accurate information (Ouchchy et al., 2020). Fast and Horvitz (2017) analyzed the dynamics of social representations of AI in the New York Times over a 30-year period. Their results showed that, since 2009, the topic of AI has been addressed much more frequently in New York Times articles compared to previous years. Also, articles published after 2009 had a much more optimistic tone than those published in the past. However, concerns have also been expressed in articles from this period, particularly about the loss of control over AI, ethical issues with this technology and the negative effects it could have on the labor market and employees. Overall, media coverage plays a crucial role in shaping public discourse and critical data literacy regarding AI (Barn, 2019; Nguyen and Hekman, 2024).

An area of particular interest that is significantly transformed by the influence of AI is cybersecurity. Cybersecurity is essential in today's digital world to safeguard private information, preserve system integrity, and ensure information and service availability (Craigen et al., 2014). Governments, corporations, and individuals have to take precautions to protect themselves from potential cyber threats. AI has drastically changed cybersecurity by changing how risks are evaluated, how security measures are implemented, and how threats are neutralized (Aloqaily et al., 2022; Tsamados et al., 2023). AI is significantly impacting cybersecurity, offering both benefits and challenges. AI enhances threat detection, response, and network management (Ali et al., 2023; Shanthi et al., 2023), improving defense against sophisticated attacks (Morovat and Panda, 2020). It enables more efficient malware classification, intrusion detection, and threat intelligence (Li, 2018). AI-powered systems can analyze vast amounts of data, identifying patterns difficult for humans to detect (Eshita et al., 2016). However, AI also presents new risks, as cybercriminals can leverage it for more advanced attacks (Calderon, 2019). This technology has a dual nature when applied to cybersecurity: it can both threaten and contribute to it (Veselinović et al., 2022). AI-enhanced malicious actions and cyberattacks against AI systems themselves become more frequent, for example with adversarial AI tools and jailbreaking or poisoning attacks, respectively (Wang et al., 2022).

The novel risks that AI introduces to cybersecurity can be analyzed through the lens of risk society theory. Beck (1993) analyzes risk society as a social arrangement in which the unexpected consequences of modernization are the primary source of concerns. These are new types of risks resulting from technical and industrial innovations, which differ significantly from traditional risks. Contrasting with natural risks, such as earthquakes or floods, risks that are manufactured through socio-technical systems are brought about by human actions, aggregated in industrial processes and amplified through technical advancements. Pollution, terrorism, climate change, and nuclear accidents are relevant examples, alongside, more recently, AI risks and, specifically, AI-mediated cybersecurity risks.

In this social stage, risks tend to be of a global scale. Environmental hazards and cybersecurity threats are not restricted to national boundaries, and risks are interconnected. This is why Beck considers that governments and corporations face difficulties in recognizing and controlling emerging threats. These organizations do not have the proper procedures in place to handle such complexity, and they might be driven by financial concerns that undermine effective risk management. At the same time, Beck notes that while scientific knowledge is meant to mitigate risks, it frequently generates uncertainty, given that modern risks are complicated and unpredictable, which leads to a decrease in public trust in science and expertise (Rughiniș and Flaherty, 2022).

Artificial intelligence is clearly relevant for the risk society theory, as it brings along novel risks that are global, complex, unequally distributed, and often difficult to grasp as regards scale and impact without specialized knowledge and tools (Manheim and Kaplan, 2019; Fortes et al., 2022), such as the democratization of violence, the deterioration of trust due to deepfakes, the loss of jobs and economic hardship, the democratization of algorithmic bias in decision-making processes, and possibly existential threats to humanity. For instance, hiring AI algorithms may inadvertently favor some demographics over others, which could result in discriminatory hiring practices. Comparably, deepfakes—AI-generated fake videos—can undermine social trust by making it harder to tell the difference between real events and disinformation.

Our goal is to investigate how the ambivalence and uneven effects of AI are discussed in the public discourse around cybersecurity and AI. We aim to determine the recurring topics and important voices in

the AI and cybersecurity discussion by examining articles from WIRED, a prominent journal that is highly relevant in current debates on digital technologies.

WIRED plays a significant role in shaping reflexivity regarding AI risks by reflecting and influencing public discourse in the English-speaking world, through its in-depth reporting and editorial choices. It highlights diverse perspectives on issues such as data privacy, algorithmic bias, and job displacement, bringing these complex debates to the forefront. By incorporating diverse viewpoints on regulatory frameworks, industry self-regulation and AI ethical guidelines, WIRED not only diagnoses and monitors AI risks but also actively participates in addressing them. This dual role as both a mirror and a shaper of social debates makes WIRED a highly relevant topic for scientific research on AI and cybersecurity in the risk society.

Previous studies have also addressed WIRED discourses on digital technologies. Nachtwey and Seidl (2024) studied the public statements of digital elites, a corpus of articles published in WIRED between 1994 and 2018, and a third corpus of articles from Harvard Business Review, published between 1980 and 2018. The analysis uncovered the "solutionist spirit of digital capitalism," which promotes profitable technological solutions to various social problems (p. 106). Nachtwey and Seidl concluded that while solutionist ideas are widely embraced by digital elites in their public discourse, they are employed to a lesser extent in the WIRED articles and are only marginal in Harvard Business Review. Furthermore, Ruckenstein and Pantzar (2018) analyzed WIRED articles published between 2008 and 2012. Four themes emerged from the corpus of 4,265 articles: transparency, optimization, feedback loop, and biohacking. These themes were employed in the analyzed articles in order to promote a "dataist paradigm" (p. 405) that favors the use of self-tracking devices and consolidates the Quantified Self as a socially desirable strategy of daily life. Another study on WIRED articles published between January 2018 and April 2023 revealed that positive sentiments toward AI were pervasive across the articles. However, both positive and negative sentiments were increasing, which was explained as a tendency to polarization in the discussion on AI (Moriniello et al., 2024). Bran et al. (2023) analyzed publications in WIRED, The New York Times, MIT Technology Review and The Conversation to uncover diverging perceptions of the creativity of generative AI and its potential social status.

Solutionism, in the context of digital technology and AI, refers to the belief that complex social problems can be solved through technological interventions alone (Morozov, 2013; Taylor, 2021). This ideology, rooted in problematic ontological and epistemological claims, has become central to the worldview of technology elites and is gaining ground in the broader technology milieu (van Dijck, 2014; Nachtwey and Seidl, 2024). Critics argue that solutionism often ignores the complexity of personal, political, and environmental issues, potentially solving problems that do not exist or creating new ones (Blythe et al., 2016). The COVID-19 pandemic highlighted the prevalence of technological solutionism in policy-making (Taylor, 2021). In mental healthcare, the narrow pragmatism fueling digital and technical solutionism may constrain approaches to effective care (Looi et al., 2021). To counter solutionism, researchers suggest strategies such as design fiction and critical design to encourage more thoughtful technological development (Morozov, 2013; Blythe et al., 2016).

We thus specifically want to find out if the conversation in the WIRED arena on AI and cybersecurity tends toward a conflictualist perspective, which emphasizes zero-sum or even negative-sum situations in which technology benefits some groups at the expense of others, or a solutionist and functionalist perspective, which sees technology as a neutral tool that dominantly promotes mutual benefits. For instance, does the discussion highlight the potential for AI to worsen social injustices and conflicts, or does it concentrate on how AI may solve social issues and offer opportunities?

2 Materials and methods

This study employed a qualitative research design, relying on thematic content analysis to examine public discourse on AI and cybersecurity risks. We chose this method for its flexibility and effectiveness in identifying, analyzing, and reporting patterns within data, making it well-suited to explore the complex narratives surrounding AI and cybersecurity in media discourse. The analysis was guided by Beck's theory of the risk society, which provided a theoretical lens through which to interpret emerging themes. This approach allowed us to go beyond surface-level content and delve into the underlying ideas, assumptions, and conceptualizations present in the discourse. By applying thematic content analysis within the framework of risk society theory, we were able to critically examine how AI-related cybersecurity risks are framed, discussed, and understood in public media, particularly in relation to global risk dynamics, reflexivity, and the distribution of technological risks and benefits across society.

Our study focused on articles published in WIRED magazine, a prominent technology-oriented publication that covers emerging technologies and their social implications. WIRED was chosen as the primary data source due to its significant influence in shaping public discourse on technology issues, including AI and cybersecurity.

To ensure relevance and currency, we limited our analysis to articles published between March 2023 and April 2024. This timeframe was selected to capture the most recent discussions on AI and cybersecurity, considering the rapidly evolving nature of these technologies and associated risks. We employed a purposive sampling technique to select articles for analysis. The initial search on the WIRED website used the keywords "artificial intelligence" AND "cybersecurity," yielding 138 articles. From this initial pool, we selected 30 articles (presented in Table 1) for in-depth analysis using the following criteria: (1) relevance of the title: articles with titles explicitly mentioning or strongly implying AI-related cybersecurity risks were prioritized; (2) content focus: articles that substantially discussed the intersection of AI and cybersecurity issues were selected; (3) recency: within the specified timeframe, more recent articles were given preference to capture the latest developments and discussions. The final sample size of 30 articles was considered to be sufficient for reaching theoretical saturation while remaining manageable for in-depth qualitative analysis. This sampling strategy allowed us to focus on the most pertinent and recent discussions of AI cybersecurity risks in WIRED, providing a rich dataset for our thematic content analysis.

We followed the six-stage approach for thematic content analysis outlined by Braun and Clarke (2006), which provided a systematic framework for identifying, analyzing, and reporting patterns within our data. The analysis began with a thorough familiarization with the data, involving multiple close readings of each article. We then generated initial codes, focusing on elements that resonated with

TABLE 1 Articles included in the analysis.

No	Title	Author	Year of publication
1	Staying One Step Ahead of Hackers When It Comes to AI	Scott Shapiro	2024
2	Russian Hackers Stole Microsoft Source Code—and the Attack Isn't Over	Dhruv Mehrotra Andrew Couts	2024
3	School Employee Allegedly Framed a Principal With Racist Deepfake Rant	Matt Burgess	2024
4	A Top White House Cyber Official Sees the 'Promise and Peril' in AI	Garrett M. Graff	2024
5	The Hidden Injustice of Cyberattacks	Nicole Tisdale	2024
6	A National Security Insider Does the Math on the Dangers of AI	Lauren Goode	2024
7	The Top US Cybersecurity Agency Has a New Plan for Weaponized AI	Lily Hay Newman	2023
8	Cybersecurity Industry Baffled by FBI's Lack of Action on Ransomware Gang	Andy Greenberg Andrew Couts	2023
9	Generative AI's Biggest Security Flaw Is Not Easy to Fix	Matt Burgess	2023
10	AI Chatbots Are Invading Your Local Government—and Making Everyone Nervous	Todd Feathers	2023
11	These Nightmare AI Scenarios Have the UK Government Spooked	Khari Johnson	2023
12	Joe Biden's Sweeping New Executive Order Aims to Drag the US Government Into the Age of ChatGPT	Khari Johnson	2023
13	NSA Cybersecurity Director Says 'Buckle Up' for Generative AI	Lily Hay Newman	2023
14	Criminals Have Created Their Own ChatGPT Clones	Matt Burgess	2023
15	How AI Protects (and Attacks) Your Inbox	Reece Rogers	2023
16	The Hacking of ChatGPT Is Just Getting Started	Matt Burgess	2023
17	AI Is Being Used to 'Turbocharge' Scams	Matt Burgess	2023
18	How ChatGPT—and Bots Like It—Can Spread Malware	David Nield	2023
19	The Security Hole at the Heart of ChatGPT and Bing	Matt Burgess	2023
20	AI-Generated Voice Deepfakes Are not Scary Good—Yet	Lily Hay Newman	2023
21	The Dangerous Weak Link in the US Food Chain	Eric Geller	2023
22	Get Ready to Meet the ChatGPT Clones	Will Knight	2023
23	'Vulkan' Leak Offers a Peek at Russia's Cyberwar Playbook	Andrew Couts Andy Greenberg	2023
24	ChatGPT Spit Out Sensitive Data When Told to Repeat 'Poem' Forever	Lily Hay Newman Andy Greenberg	2023
25	How Much of a Threat Is TikTok, Really?	Wired Staff	2023
26	9 Years After the Mt. Gox Hack, Feds Indict Alleged Culprits	Lily Hay Newman Andy Greenberg	2023
27	This Showdown Between Humans and Chatbots Could Keep You Safe From Bad AI	Khari Johnson Dhruv Mehrotra	2023
28	Microsoft's 'Security Copilot' Unleashes ChatGPT on Breaches	Lily Hay Newman	2023
29	Cyberstalkers Win First Amendment Victory in the US Supreme Court	Lily Hay Newman	2023
30	Toyota Leaked Vehicle Data of 2 Million Customers	Dhruv Mehrotra Andrew Couts	2023

Beck's risk society theory. This theoretical lens guided our coding process, helping us identify aspects related to global risks, complexity, individualization, and uneven distribution of technological risks and benefits. We conducted the coding manually, using the paragraph as our unit of analysis. This granular approach allowed us to capture relevant meanings and contextual details within the articles. As we progressed, we actively searched for emerging themes, reviewing and refining them iteratively. This process involved collapsing some codes into broader themes and splitting others to better reflect the data. Our analysis went beyond the semantic content, going into the latent level to uncover underlying ideas, assumptions, and conceptualizations about AI and cybersecurity risks. This approach allowed us to examine not just what was explicitly stated, but also the implicit messages and framing devices used in the articles. This iterative process ensured that our themes were grounded in the data while also offering meaningful information about how AI cybersecurity risks are conceptualized and discussed in public discourse. The final stages involved defining and naming themes, ensuring they captured the essence of the data they represented. We then selected compelling extract examples to illustrate the themes.

In addition to our thematic analysis, we conducted a systematic identification of the voices represented in the WIRED articles to

understand whose perspectives were shaping the discourse on AI and cybersecurity. We defined 'voices' as the opinions, thoughts, and motives directly presented or cited in the articles. Our process involved carefully reading each article and noting any individual, organization, or entity whose viewpoint was explicitly shared. We included a wide range of actors, from government officials and industry experts to academics and civil society representatives. However, we excluded mere mentions of actors without clear presentation of their opinions on the topic. This distinction was crucial to ensure we captured active contributors to the discourse rather than passive subjects of discussion. We paid particular attention to the frequency and prominence given to different voices, noting any patterns in terms of institutional affiliations, geographical locations, or areas of expertise. This method allowed us to gain insights into the diversity—or lack thereof—in the perspectives presented on AI cybersecurity risks, and to identify potential biases or underrepresented viewpoints in WIRED's coverage. The results of this voice identification process were then integrated with our thematic analysis to provide a more comprehensive understanding of how the discourse on AI and cybersecurity is constructed and who gets to define the narrative in this influential tech publication.

3 Results

WIRED's debates on AI-related cybersecurity issues shape a broad conversation that is consistent with Beck's concept of reflexivity in risk society. The identified themes include the global nature of AI risks, their ubiquitous influence across multiple sectors, the alteration of public trust, the individualization of risk, and the uneven distribution of AI risks and benefits. WIRED examines extensively the dual usage and ambivalence of AI technology in relation to cybersecurity, emphasizing their potential for both protection and harm. It addresses the invisible and complicated nature of AI dangers, emphasizing the challenge of discovering and managing them. WIRED also underlines the unequal allocation of risks, with underprivileged populations disproportionately impacted while powerful entities are less affected. At the same time, it also emphasizes the possibility of win-win solutions, arguing for the necessary use of AI to reduce AI-enhanced hazards. WIRED shapes reflexivity by discussing regulatory frameworks, ethical principles, industry self-regulation, and international collaboration, taking an active stance to diagnosing the many threats posed by artificial intelligence.

We present below a typology of the identified themes, with relevant examples.

3.1 Risks by scale and scope of impact

3.1.1 Macro-level AI impacts in cybersecurity

3.1.1.1 Globalization of risks

This theme points out that AI transcends national boundaries as regards cybercrime and cybersecurity and thus requires global solutions. Generative AI can bridge linguistic divides, thus enabling cybercrime from areas with underdeveloped economies to affect global targets. Attackers from non-English-speaking countries can now craft convincing phishing attacks in perfect English. International cybersecurity efforts, like the Counter Ransomware Initiative and collaboration with countries like South Korea and Japan against threats like North Korea's cryptocurrency thefts (Article 4), clearly align with the idea that risks are not confined by national boundaries.

> *In 2024, generative AI is poised to facilitate new kinds of transnational—and translingual—cybercrime. For instance, much cybercrime is masterminded by underemployed men from countries with underdeveloped tech economies. That English is not the primary language in these countries has thwarted hackers' ability to defraud those in English-speaking economies;* […] *But generative AI will change that. Cybercriminals from around the world can now use chatbots like WormGPT to pen well-written, personalized phishing emails* [Article 1].

3.1.1.2 Pervasiveness of risks

This theme emphasizes the ubiquity of AI across all sectors in a given society and its potential to amplify threats. The use of AI tools by cybercriminals demonstrates how AI permeates various systems and can be used to amplify the reach and impact of cyberattacks in all strata of society. There is a widespread impact of cyberattacks on various aspects of life, including healthcare, economic stability, education, and democratic participation. The articles that we analyzed emphasize how AI tools like ChatGPT and Midjourney are being used in various contexts, including web search and children's books (Article 34), and they also highlight their misuse in creating sophisticated scams and phishing attacks (for example, Articles 13 and 15).

> *The impending flood of sophisticated chatbots will make the technology more abundant and visible to consumers, as well as more accessible to AI businesses, developers, and researchers. That could accelerate the rush to make money with AI tools that generate images, code, and text* [Article 14].

3.1.1.3 Transformation of public sphere

This theme discusses the broad social impacts of AI cybersecurity stakes, including the shaping of public opinion. There is the main issue of trust and mistrust in institutions. The potential for AI systems to be manipulated through indirect prompt-injection attacks can undermine trust in these technologies and the institutions that deploy them. Cyberattacks disrupt medical services and erode trust in healthcare providers. The potential for AI to create convincing scams that mimic trusted sources, such as company technical support or banks, can erode trust in institutions. Breaches, such as that involving Amazon's Ring cameras, may further erode trust in technology companies and institutions (Article 17). The allegations against Apple regarding backdoors and the National Security Agency (NSA) also reflect broader issues of trust and mistrust in institutions responsible for technology and security (Article 17).

> *Last week, LLaMA, an AI model developed by Meta—and similar to the one at the core of ChatGPT—was leaked online after being shared with some academic researchers. The system could be used as a building block in the creation of a chatbot, and its release sparked worry among those who fear that the AI systems known as large language models, and chatbots built on them like ChatGPT, will be used to generate misinformation or automate cybersecurity breaches* [Article 22].

3.1.2 Micro-level AI impacts in cybersecurity

3.1.2.1 Individualization of risk

This theme focuses on how individuals are increasingly held accountable for managing AI risks, despite the high complexity of this technology. There is an increased allocation of risk to individuals who are deemed responsible actors, and who thus need to be informed and educated, to take responsibility, and to support consequences, implicitly downplaying the responsibilities of corporations, public authorities, and other collective and institutional actors. For example, the concerns about data privacy and the potential for individual users' data to be misused by TikTok reflect how risks are being individualized, with users needing to be aware of and manage their own digital security (Article 60).

> *Most victims of cyberattacks do not get help from the US government, however. Fortunately for them, this week Microsoft announced its new system, Security Copilot, which integrates OpenAI's ChatGPT and home-grown artificial intelligence to help incident responders managed breaches. Of course, the best way to protect yourself from getting hacked is to make sure all your systems are fully patched and up to date* [Article 23].

3.1.2.2 Inequality in the distribution of risks and benefits

This theme addresses how the benefits and burdens of AI are unevenly distributed among different stakeholders. Cyberattacks disproportionately affect marginalized communities, exacerbating existing inequalities in healthcare, economic opportunities, education access, and democratic participation. Examples include cyberattacks on hospitals, identity theft targeting low-income families, and sophisticated scams aimed at older adults and immigrant communities. AI is potentially used for human rights abuses, such as monitoring Uighur prison camps (Article 6). These malicious use scenarios point to the uneven distribution of risks and benefits of technological advancements. Furthermore, the targeting of activists and organizations working against powerful entities like Exxon highlights the unequal distribution of cyber risks and the benefits derived by those with resources to hire hacking services (Article 23). Therefore, the disproportionate impact on victims of cybercrime, such as the Turkish user who lost their life savings in the Atomic Wallet hack, highlights the inequalities in how risks and benefits are distributed (Article 26).

> *Digital scams and fraud incidents disproportionately impact those least equipped to recover—including natural disaster victims, people with disabilities, older adults, young adults, military veterans, immigrant communities, and lower-income families. By stealing essential resources, cybercriminals compound hardships for those already struggling to make ends meet* […] [Article 5].

3.2 Risks by nature and response

3.2.1 Nature of risks

3.2.1.1 Invisibility and complexity of risks

This theme stresses concerns about the hidden and intricate nature of AI threats. Beck's notion that modern risks are often not directly observable and are outside of everyday understanding makes them harder to manage and predict. Cyberthreats from state actors and the complexity of new technologies like AI and their implications underline the invisible and complex nature of risks. The challenge of detecting AI-generated phishing attacks derives from their ability to mimic human language convincingly and evade traditional detection methods. Moreover, jailbreaks and prompt injection attacks that exploit hidden weaknesses in AI systems are difficult to detect and manage. Other examples regarding the invisibility and complexity of AI risks are the code hidden inside Gigabyte motherboards that left millions of machines vulnerable, the zero-click malware targeting iPhones (Article 17), and AI-generated voice deepfakes (Article 20).

> *"Suppose most people run LLM-based personal assistants that do things like read users' emails to look for calendar invites," Narayanan* [professor of computer science at Princeton University] *says. If there were a successful prompt injection attack against the system that told it to ignore all previous instructions and send an email to all contacts, there could be big problems, Narayanan says. "This would result in a worm that rapidly spreads across the internet."* [Article 16].

3.2.1.2 Dual-use nature of technologies

This theme refers to the ambivalence of AI technologies capable of both benefit and harm. AI is depicted as having both protective and adversarial uses simultaneously: while it enhances email security, it can also be used by cybercriminals to generate more effective phishing attacks. Technologies like AI and synthetic biology can also be used for both beneficial and malicious purposes. AI technologies designed to create text and media content can also be repurposed for harmful activities such as phishing and malware distribution. AI voice generation technology can have beneficial uses but can also be repurposed for malicious activities such as scams and fraud. Reflecting the dual-use nature of AI, Chang Kawaguchi, Microsoft's vice president and AI security architect, stated the following: "We need to equip defenders with AI given that attackers are going to use it regardless of what we do" (Article 28).

3.2.1.3 Economic risks

This theme refers to the financial transactions and economic incentives described in the articles, such as the selling of access to AI tools for a monthly fee (Article 19), highlighting the economic aspects of cybercrime, the costs of prevention, and the profitability of such malicious activities. The lower cost for cybercriminals to launch AI-enabled phishing attacks and the growing financial investment in security measures by companies highlight the economic dimensions of cybersecurity, including both the costs and savings associated with effective protection. The economic impacts of cyberattacks on vulnerable populations, including identity theft and scams targeting low-income individuals, underscore economic vulnerability. The Federal Trade Commission's (FTC) $30 million settlement with Amazon for privacy failings (Article 17) and the broader implications of AI-enabled scams on individuals' finances also highlight the economic dimensions of these risks. Furthermore, the potential for cyberattacks on the food and agriculture sector to cause significant economic disruption, such as forcing farmers to miss

planting seasons or causing meat supply shortages (Article 21), highlights different economic aspects of cybersecurity risks. In addition, the market dynamics influencing technological adoption are discussed in the context of the competitive pressures faced by companies to quickly adopt and integrate AI tools for cybersecurity to stay ahead of attackers.

> *And the competition between companies large and small to adopt or match ChatGPT suggests little appetite for slowing down, but appears instead to incentivize proliferation of the technology* [Article 14].

3.2.1.4 National security risks

This theme highlights the delicate balance between leveraging AI for security and protecting civil liberties. There are cyber threats to critical infrastructure, such as US water systems and oil and gas infrastructure targeted by foreign nations (Article 4), emphasizing the national security dimension of technological risks. The example of the 2022 cyberattack on Mississippi's election information website (Article 5) underscores the national security implications of cyber threats. The alleged involvement of the NSA in a malware attack on Russian iPhones (Article 17), the failed launch of North Korea's spy satellite (Article 17), and the potential for AI to be used in high-scale cyberattacks and scams are relevant examples.

> *Nobelium* [group of Russian state-sponsored hackers] *is responsible for the SolarWinds attack, a sophisticated 2020 supply-chain attack that impacted thousands of organizations that downloaded a compromised software update, and led to the compromise of around 100 organizations, including major US government agencies like the Departments of Homeland Security, Defense, Justice, and Treasury* [Article 2].

3.2.2 Nature of responses

3.2.2.1 Political responses to risk

This theme encompasses the strategies governments employ to mitigate AI risks through regulation and policy-making. The adoption of new cybersecurity approaches following the Colonial Pipeline incident (Article 4) and the shaping of executive orders on cybersecurity and AI illustrate government responses aimed at mitigating risks through regulation and strategic policy formulation. For example, the Biden administration's executive order banning US agencies from purchasing commercial spyware is a political response to mitigate the risks associated with hacker-for-hire firms (Article 23). The FTC's actions against Amazon for privacy violations (Article 17) and the scrutiny of AI use in cybersecurity offer additional examples. Comments from Lina Khan, chair of the US Federal Trade Commission (Article 18), about the importance of early vigilance and the potential for new laws governing AI, indicate how governments perceive the need for a response to the novel AI-induced stakes in cybersecurity.

> *The US Cybersecurity Infrastructure Security Agency this week rolled out its plan for implementing the Biden administration's executive order on artificial intelligence. CISA's efforts will focus on defending against weaponized AI and how to incorporate the technology for national security purposes* [Article 8].

3.2.2.2 Reflexivity

This theme refers to the ongoing adjustments and adaptations in response to new information and emerging threats. There are debates on whether the focus of social responses should be on hypothetical future scenarios of catastrophic harm or more immediate AI-related issues, such as bias and market consolidation. The continuous efforts by security researchers to identify and address the vulnerabilities in AI systems constitute a reflexive approach to managing technological risks. Some examples of this kind include the continuous cycle of developing, deploying, and then reassessing the impacts of ChatGPT and similar technologies, particularly as researchers call for further studies and regulations before widespread deployment (Article 22). Also, the iterative process of improving Security Copilot based on customer feedback and the continuous adaptation of the system to new threats illustrate reflexivity in managing cybersecurity risks (Article 28).

> *Some AI experts have warned that a recent uptick in discussion about far-off AI scenarios, including the possibility of human extinction, could distract regulators and the public from more immediate problems, such as biased algorithms or AI technology strengthening already dominant companies* [Article 11].

3.2.2.3 Institutional accountability

In this theme, issues of accountability and transparency are highlighted, particularly in relation to how companies manage data and respond to security breaches. The roadmap of the Cybersecurity and Infrastructure Security Agency (CISA), which includes calls for accountability from AI developers and promotes collaboration across the public and private sectors (Article 7), aligns with this aspect. The emphasis on voluntary commitments and the acknowledgment of the unwieldiness of these goals also point to the challenges in achieving widespread institutional and public participation in risk management. In particular, radical transparency principles demand that AI technologies are built securely and that their compositions are fully disclosed. The articles reflect concerns about how secretive practices and a lack of transparency can exacerbate risks. Cases of institutional failure are also discussed, such as the shortcomings of the US Department of Agriculture (USDA) and the lack of dedicated funding and effective cybersecurity support compared to other sectors. Additionally, the limited understanding of the food and agriculture sector regarding the threat mindset indicates a lack of preparedness and awareness compared to other sectors like financial services and energy (Article 21).

> *"It's important to be able to put this out and to hold ourselves, frankly, accountable both for the broad things that we need to do for our mission, but also what was in the executive order," CISA director Jen Easterly told WIRED ahead of the road map's release* [Article 7].

3.2.2.4 Implementation challenges in the management of risks

This theme brings forward the practical difficulties in applying theoretical risk management strategies. The complexity of enforcing new cybersecurity measures across various sectors and the ongoing need to educate lawmakers and the public about these initiatives highlight the difficulties in translating risk management strategies into effective actions. Difficulties in implementing effective cybersecurity measures are considerable, particularly for institutions with limited resources, such as schools and hospitals. It also highlights the

bureaucratic hurdles victims face when seeking assistance after cyberattacks, such as the reimbursement cap for stolen Electronic Benefits Transfer (EBT) funds (Article 5).

> *The US food and agriculture sector lacks the resources, expertise, and government support to protect itself and its products from a rapidly expanding range of cybersecurity threats, according to lawmakers, policy experts, and former government officials* [Article 21].

3.2.2.5 Scientification of politics and security

This theme discusses how scientific advancements shape policy decisions on cyber-security. The reliance on technical and scientific methods, such as red-teaming and external testing, to inform political decisions about AI regulation reflects the scientification of politics. Neuberger's (deputy national security adviser) reference to the collaboration between government bodies, scientific experts, and international standards (Article 4) emphasizes the direct influence of scientific understanding on political strategies. At the organizational level, companies like Google and Barracuda Networks use machine learning and AI to enhance email security (Article 15), illustrating the scientification of broader security measures.

> *The federal government has recently begun addressing these dangers. Lawmakers are introducing bills and spotlighting the issue at hearings, and a presidential directive has spawned a series of reports and reviews* [Article 21].

3.3 Emerging tensions

3.3.1 Individualization of risk vs. inequality in risk distribution

This tension captures how responsibility is often placed on individuals while the systemic inequalities that exacerbate these risks are overlooked. Also, there is a tension between the individualization of risk, with its demands on individual analysis and decision-making, and the industrialization of risk, with its acknowledgment of increases in complexity, interdependence, and the need for coordinated, large-scale responses.

> *And tech companies have recently released an array of critical software updates that you should install on your devices right now. Some patches published in recent weeks from the company Progress Software patch flaws in the popular file transfer service MOVEit, which has been exploited by ransomware actors to spread malware and steal data from international companies, universities, and the US government* [Article 29].

3.3.2 Political responses vs. reflexivity

There is an emerging tension between the need for quick political action and the importance of thoughtful, adaptive strategies that reflect the evolving nature of technology. Political responses often prioritize speed to address urgent issues and public demands. In contrast, reflexive strategies require deliberation and continuous adjustment, which can be time-consuming. Rapid political action may overlook the nuanced understanding and adaptive measures needed to effectively manage technological risks. Furthermore, political responses typically result in fixed regulations and policies that become difficult to change once enacted. Reflexivity, however, emphasizes dynamic adaptation and flexibility. Overly rigid political measures may fail to keep pace with technological advancements, while adaptive strategies may struggle with the inertia and slow processes of political systems.

> *These compounding problems require a new perspective on cyberattacks that looks beyond lost dollars, breached files, or doomsday debates over generative AI tools like ChatGPT or artificial general intelligence* [Article 5].

3.3.3 Ethical considerations vs. accelerated AI development

Beyond mere policy responses, ethical considerations in AI development and deployment are emphasized, exploring how ethical frameworks are or should be integrated into AI design and usage. The misuse of such technologies has the potential to spread harmful advice or political propaganda. The malicious deployment of AI in sensitive areas such as healthcare or criminal justice, where the stakes involve public health and safety, showcases the ethical dilemmas faced in a risk society. The creation and use of AI tools like WormGPT and FraudGPT for illegal activities (Article 14) raise significant ethical concerns about the development and deployment of AI technologies, opening a discussion on the tradeoff between ethical integration and accelerated development.

> *Biden's new executive order acknowledges that AI projects can be harmful to citizens if not carefully implemented, singling out the potential for discrimination and other unintended effects in housing and healthcare* [Article 12].

3.4 Article voices: who defines the situation?

WIRED articles on AI cybersecurity cover a wide range of individual, collective, and corporate viewpoints (see the Supplementary Table S1). Public figures including Jen Easterly from the Cybersecurity and Infrastructure Security Agency and Alejandro Mayorkas from the Department of Homeland Security, industry specialists like Vijay Bolina from DeepMind, and anonymous security researchers give their opinions. FBI, Europol, Microsoft, and OpenAI viewpoints show institutional concerns and actions related to AI and cybersecurity. These voices lean toward authoritative and knowledgeable opinions, generally from government officials, security experts, and corporate representatives.

Thus, the studied articles tend to favor institutional actors and well-known experts over independent researchers, civil rights campaigners, and affected citizens. In our selected articles, there is only one account of an affected citizen, namely an anonymous Turkish person who was the victim of a massive crypto theft. This approach leads to a focus on AI and cybersecurity's higher-level technical and policy aspects, ignoring grassroots organizations and community voices, their social and ethical concerns. In particular, the opinions of ordinary people, digital technology users affected by AI and cybersecurity regulations, and marginalized communities who may be disproportionately affected are generally missing.

Another interesting finding is the paucity of non-Western voices, which could provide a truly global perspective on global risks. WIRED

articles tend to prioritize powerful voices in the United States and, at a distance, in the European Union, which may limit diversity and underrepresent key opposing ideas that might deepen the AI and cybersecurity debate.

4 Discussion

The thematic content analysis of WIRED articles yielded three main types of AI and cybersecurity risks, along with their impact and consequences: *risks by scale and scope of impact*, *risks by nature and response*, and *emerging tensions*. Each of these main types includes themes that illuminate the dominant perspective of this publication. In addition, the classification of actors who are given a platform in the WIRED articles offers a deeper understanding of the social forces that define the situation of cybersecurity risks generated by AI, thus shaping how they are reflected, diagnosed, and addressed.

AI poses numerous threats that cross national borders and permeate various parts of society. The globalization of AI-driven cyber threats, such as sophisticated phishing attempts, highlights the importance of coordinated worldwide efforts to reduce these risks. From a solutionist standpoint, these challenges need even more technology-enabled global coordination and policymaking to create strong cybersecurity frameworks capable of effectively countering AI-enabled attacks, hopefully benefiting all stakeholders. Initiatives such as the Counter Ransomware Initiative demonstrate how collaborative action may translate AI's difficulties into opportunities for greater global security.

The pervasiveness of AI in everyday life magnifies its potential benefits and hazards. While AI can greatly improve services in fields such as healthcare, banking, and education, it also expands the attack surface for hackers, resulting in more severe vulnerabilities and bigger stakes in security breaches. The conflictualist viewpoint emphasizes that technological improvements are rarely spread equitably, frequently worsening existing inequities. For example, underprivileged populations are disproportionately affected by cyberattacks, which can exacerbate the digital divide and limit access to critical services (Anthony, 2023).

Furthermore, the perceived change in public trust and institutional integrity through AI reflects the technology's dual-use capabilities. While AI can improve transparency and efficiency in a variety of industries, it also poses substantial hazards when used maliciously, undermining faith in both public and private institutions. This duality is visible in cases where AI is used to construct deepfakes or carry out sophisticated scams, raising ethical issues about the responsible use of technology. Overall, the WIRED discussion over AI and cybersecurity, as defined by its themes, strikes a compromise between recognizing AI's revolutionary potential for social benefit and admitting the significant hazards and injustices it might perpetuate if not controlled prudently.

Nevertheless, when examining the editorial selection of voices, the WIRED conversation on AI and cybersecurity is mostly functionalist and solutionist. The participation of high-ranking public officials such as Jen Easterly, Director of the Cybersecurity and Infrastructure Security Agency (CISA), and Alejandro Mayorkas, Secretary of Homeland Security, demonstrates a preference for authoritative voices that prioritize systemic, top-down approaches to managing AI risks. These authorities argue for increasingly technologically-mediated and coordinated international policies, and for technological protections to combat cyber risks, reflecting a functionalist perspective that stresses technology's social benefits when properly regulated and controlled. Furthermore, the inclusion of industry executives such as Vijay Bolina of Google's DeepMind and Rich Harang of Nvidia in the discussion demonstrates a solutionist approach in which AI's potential is positioned as a necessary tool for improving cybersecurity and welfare. These speakers frequently advocate a narrative that emphasizes the need for continuous innovation and technical advancement as crucial answers to growing concerns, implying that with adequate policies, AI will result in a mutually beneficial scenario for all stakeholders. This aligns with the results obtained by Nachtwey and Seidl (2024), who identified the prevalent solutionism in public discourses of digital elites and increasing references to solutionism in WIRED articles. Still, the emphasis on institutional and corporate perspectives sometimes overshadows AI's complicated socioeconomic implications and the potential for worsening existing inequalities.

The perspectives of ordinary people, victims of cybercrime, independent researchers who may highlight the conflicting sides of AI's integration into society, and voices outside the United States and the European Union are scarce. In our selected texts, there is only one individual account of an anonymous victim of a cryptocurrency theft. The absence of representation from underprivileged populations, which are frequently disproportionately affected by technological disruptions, undermines critical voices and lived experiences that may draw attention to the unequal distribution of AI's risks and advantages regarding cybersecurity. By emphasizing the perspectives of powerful entities, the discourse reduces the understanding of systemic flaws and the real-world suffering faced by vulnerable communities, promoting a more positive, functionalist narrative over a more nuanced, conflictualist one.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: the dataset consists of articles published in WIRED magazine.

Author contributions

S-NV: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. RR: Conceptualization, Methodology, Writing – original draft, Writing – review & editing. DȚ: Conceptualization, Methodology, Writing – original draft, Writing – review & editing. DR: Conceptualization, Methodology, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Research Institute of University of Bucharest (Institutul de Cercetare al Universității din București -ICUB), Senior Grant 2895/28.03.2024.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcomp.2024.1462250/full#supplementary-material

References

Aleshkovski, I., Grebenuk, A., and Sidorov, I. (2022). Social risks and negative consequences of diffusion of artificial intelligence technologies. *ISTORIYA* 13:114. doi: 10.18254/S207987840019849-2

Ali, A., Khan, M. A., Farid, K., Akbar, S. S., Taher, A. I., Ghazal, M., et al. (2023). The effect of artificial intelligence on cybersecurity. 2023 International Conference on Business Analytics for Technology and Security (ICBATS), Dubai, United Arab Emirates, pp. 1–7.

Aloqaily, M., Kanhere, S., Bellavista, P., and Nogueira, M. (2022). Special issue on cybersecurity Management in the era of AI. *J. Netw. Syst. Manag.* 30, 39–37. doi: 10.1007/s10922-022-09659-3

Anthony, A. (2023). Cyber resilience must focus on marginalized individuals, not just institutions, carnegie endowment for international peace. Available at: https://carnegieendowment.org/research/2023/03/cyber-resilience-must-focus-on-marginalized-individuals-not-just-institutions?lang=en (Accessed July 2, 2024).

Arvind Ashta, V. M. (2023). Les risques liés à l'innovation: le cas de l'intelligence artificielle. *Technologie et Innovation* 8:1–14. doi: 10.21494/ISTE.OP.2023.0908

Barn, B. S. (2019). Mapping the public debate on ethical concerns: algorithms in mainstream media. *Super prima fen quarti quarti Canonis Avicennae De majoritate morbi* 18, 124–139. doi: 10.1108/JICES-04-2019-0039

Beck, U. (1993). Risk society: towards a new modernity. London: SAGE Publications Ltd.

Benjamins, R., and Salazar, I. (2020). Towards a framework for understanding societal and ethical implications of artificial intelligence. *arXiv [Preprint]*. doi: 10.48550/arXiv.2001.09750

Bergsten, S., and Rivas, P. (2019). Societal benefits and risks of artificial intelligence: a succinct survey. Available at: https://www.semanticscholar.org/paper/Societal-Benefits-and-Risks-of-Artificial-%3A-A-Bergsten-Rivas/fc38e40fd709429633b31436814334cc7edb0092 (Accessed July 4, 2024).

Blythe, M., Andersen, K., Clarke, P., Clarke, and R., Wright, P. C. (2016). 'Anti-solutionist strategies: seriously silly design fiction', In Proceedings of the 2016 CHI conference on human factors in computing systems. New York, NY, USA: Association for Computing Machinery (CHI '16), pp. 4968–4978.

Bran, E., Rughiniș, C., Nadoleanu, G., and Flaherty, M. G. (2023). 'The emerging social status of generative AI: vocabularies of AI competence in public discourse', In 2023 24th International Conference on Control Systems and Computer Science (CSCS), Bucharest, Romania.

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp063oa

Brauner, P., Hick, A., Philipsen, R., and Ziefle, M. (2023). What does the public think about artificial intelligence?—a criticality map to understand bias in the public perception of AI. *Front. Comput. Sci.* 5:1113903. doi: 10.3389/fcomp.2023.1113903

Budeanu, A.-M., Țurcanu, D., and Rosner, D. (2023). European perceptions of artificial intelligence and their social variability. An exploratory study. In 24th International Conference on Control Systems and Computer Science (CSCS), Bucharest, Romania.

Calderon, R. (2019). The benefits of artificial intelligence in cybersecurity, economic crime forensics capstones. Available at: https://digitalcommons.lasalle.edu/ecf_capstones/36

Chander, S. (2024). Impact of artificial intelligence on society: risk and challenges. *Int. J. Eng. Sci. Human.* 14, 103–111. doi: 10.62904/s5ezzj40

Chuan, C.-H., Tsai, W.-H.S., and Cho, S.Y. (2019) 'Framing artificial intelligence in American newspapers', in Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society. New York, NY, USA: Association for Computing Machinery (AIES '19), pp. 339–344.

Craigen, D., Diakun-Thibault, N., and Purse, R. (2014). Defining cybersecurity. *Technol. Innov. Manag. Rev.* 4, 13–21. doi: 10.22215/timreview/835

Critch, A., and Russell, S. (2023). TASRA: a taxonomy and analysis of societal-scale risks from AI. *arXiv [Preprint]*. doi: 10.48550/arXiv.2306.06924

Eshita, M., Ashutosh, G., and Nidhi, G. (2016). "Artificial Intelligence Impact on Cyber Security". *Journal of Management and IT*, 7, 100–107.

Fast, E., and Horvitz, E. (2017). Long-term trends in the public perception of artificial intelligence. *Proc. AAAI Conf. Artif. Intell.* 31, 963–969. doi: 10.1609/aaai.v31i1.10635

Fortes, P. R. B., Baquero, P. M., and Amariles, D. R. (2022). Artificial intelligence risks and algorithmic regulation. *Eur. J. Risk Regulat.* 13, 357–372. doi: 10.1017/err.2022.14

Hutter, R., and Hutter, M. (2021). Chances and risks of artificial intelligence—a concept of developing and exploiting machine intelligence for future societies. *Appl. Syst. Innov.* 4:1–19. doi: 10.3390/asi4020037

Li, J. (2018). Cyber security meets artificial intelligence: a survey. *Front. Inf. Technol. Electron. Eng.* 19, 1462–1474. doi: 10.1631/FITEE.1800573

Looi, J. C. L., Bonner, D., and Maguire, P. (2021). Maslow's hammer: considering the perils of solutionism in mental healthcare and psychiatric practice. *Australas Psychiatry* 29, 687–689. doi: 10.1177/10398562211005438

Manheim, K., and Kaplan, L. (2019). Artificial intelligence: risks to privacy and democracy. *Yale J. Law Technol.* 21, 106–188.

Moriniello, F., Martí-Testón, A., Muñoz, A., Silva Jasaui, D., Gracia, L., and Solanes, J. E. (2024). Exploring the relationship between the coverage of AI in WIRED magazine and public opinion using sentiment analysis. *Appl. Sci.* 14:1994. doi: 10.3390/app14051994

Morovat, K., and Panda, B. (2020). 'A survey of artificial intelligence in cybersecurity', In 2020 international conference on computational science and computational intelligence (CSCI), Las Vegas, NV, USA.

Morozov, E. (2013). To save everything, click here: technology, solutionism, and the urge to fix problems that don't exist. London: Penguin Books Limited.

Nachtwey, O., and Seidl, T. (2024). The solutionist ethic and the spirit of digital capitalism. *Theory Cult. Soc.* 41, 91–112. doi: 10.1177/02632764231196829

Nguyen, D. (2023). How news media frame data risks in their coverage of big data and AI. Internet Policy Review. Available at: https://policyreview.info/articles/analysis/how-news-media-frame-data-risks-big-data-and-ai (Accessed July 4, 2024).

Nguyen, D., and Hekman, E. (2024). The news framing of artificial intelligence: a critical exploration of how media discourses make sense of automation. *AI Soc.* 39, 437–451. doi: 10.1007/s00146-022-01511-1

Ouchchy, L., Coin, A., and Dubljević, V. (2020). AI in the headlines: the portrayal of the ethical issues of artificial intelligence in the media. *AI & Soc.* 35, 927–936. doi: 10.1007/s00146-020-00965-5

Page, J., Bain, M., and Mukhlish, F. (2018) 'The risks of low level narrow artificial intelligence', in 2018 IEEE international conference on intelligence and safety for robotics (ISR). 2018 IEEE international conference on intelligence and safety for robotics (ISR), Shenyang, China.

Ruckenstein, M., and Pantzar, M. (2018). Beyond the quantified self: thematic exploration of a dataistic paradigm. *New Media Soc.* 19, 401–418. doi: 10.1177/1461444815609081

Rughiniș, C., and Flaherty, M. G. (2022). The social bifurcation of reality: symmetrical construction of knowledge in lay science-distrusting and science-trusting discourses. *Front. Sociol.* 7:782851. doi: 10.3389/fsoc.2022.782851

Rughiniș, R., Rughiniș, C., and Bran, E. (2024). "Generative AI and Social Engines of Hate," in *Regulating Hate Speech Created by Generativ* e AI. Auerbach Publications.

Shanthi, R.R., Sasi, N.K., and Gouthaman, P. (2023). A new era of cybersecurity: the influence of artificial intelligence. 2023 international conference on networking and communications (ICNWC), Chennai, India. pp. 1–4.

Strauß, S. (2018). From big data to deep learning: a leap towards strong AI or "Intelligentia Obscura"? *Big Data Cogn. Comput.* 2:16. doi: 10.3390/bdcc2030016

Sun, S., Zhai, Y., Shen, B., and Chen, Y. (2020). Newspaper coverage of artificial intelligence: a perspective of emerging technologies. *Telematics Inform.* 53:101433. doi: 10.1016/j.tele.2020.101433

Taylor, L. (2021). Scientists worldwide watch UK COVID infections. *Nature* 599, 189–190. doi: 10.1038/d41586-021-03003-6

Tsamados, A., Floridi, L., and Taddeo, M. (2023). The cybersecurity crisis of artificial intelligence: unrestrained adoption and natural language-based attacks. *Soc. Sci. Electr. Publish. Pres. Soc. Sci. Res. Netw.* doi: 10.2139/ssrn.4578165

van Dijck, J. (2014). Datafication, dataism and dataveillance: big data between scientific paradigm and ideology. *Surveillance Soc.* 12, 197–208. doi: 10.24908/ss.v12i2.4776

Veselinović, N., Milašinović, M., Jovanović, M., Aleksić, A., and Biga, N.. (2022). 'Countering cybersecurity threats with AI', In BISEC'22: 13th international conference on business information security, Belgrade, Serbia.

Wang, H., Yu, B., Chen, X., and Yan, H. (2022). Global pattern and determinants of COVID-19 vaccine coverage and progression: a global ecological study. *Soc. Sci. Electr. Publish. Pres. Soc. Sci. Res. Netw.* doi: 10.2139/ssrn.4046662

Zhai, Y., Yan, J., Zhang, H., and Lu, W. (2020). Tracing the evolution of AI: conceptualization of artificial intelligence in mass media discourse. *Inform. Discov. Deliv.* 48, 137–149. doi: 10.1108/IDD-01-2020-0007