Media and Communication
2026 • Volume 14 • Article 11242
https://doi.org/10.17645/mac.11242

cogitatio

ARTICLE

Open Access Journal

Governing Health Risk Communication in the Age of AI: Approaches from Brazil and Germany

Ana Carolina Monari [1,†] and Tabea Ott [2,3,4,†]

[1] Artificial Intelligence Laboratory (Recod.ai), University of Campinas, Brazil
[2] Institute of Ethics and Law in Medicine, University of Vienna, Austria
[3] Institute of Systematic Theology and the Study of Religions, University of Vienna, Austria
[4] Faculty of Humanities, Social Sciences, and Theology, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

Correspondence: Tabea Ott (tabea.ott@univie.ac.at)

[†] These authors contributed equally to this work

Submitted: 29 August 2025 **Accepted:** 9 December 2025 **Published:** 5 February 2026

Issue: This article is part of the issue "Communicating Risk, Trust, and Resilience Among Diverse and Marginalized Populations" edited by Ian Somerville (University of Leicester) and Jenny Zhengye Hou (Queensland University of Technology), fully open access at https://doi.org/10.17645/mac.i507

Abstract

Misinformation, disinformation, and malinformation pose a challenge to health risk communication. These information disorders can be amplified and altered by AI. In this article, we compare governance approaches from Germany and Brazil to understand how these two democracies address the role of AI in both exacerbating and mitigating health information disorders. Drawing on document analysis of hard and soft law governance approaches, the study shows that both countries frame AI as an ambivalent tool: While AI enables large-scale production and spread of misleading content, it also provides mechanisms for detection and monitoring of this content, and allows tailored communication. In Germany, information disorders are addressed through a multi-level governance strategy that combines national regulation with international frameworks. At the same time, independent fact-checking organizations and individual actors play a complementary role by monitoring public discourse, verifying claims, and fostering media literacy. In Brazil, efforts include legislative proposals on AI, the National AI Plan, and fact-checking services, though implementation is complicated by political and economic dynamics. Across both contexts, governance measures emphasize individual and organizational responsibilities, particularly through platform regulation, transparency, and digital literacy, but often underplay broader societal and interactional factors such as political structures, trust, and post-truth dynamics.

Keywords

disinformation; ethics; governance; infodemics; malinformation; misinformation; politics; risk communication; vulnerable communities

1. Introduction

Health risk communication plays an important role in protecting and promoting public health. It involves conveying clear, accurate, and timely information to help individuals and communities make informed decisions about their well-being. However, in the digital age, this task has become increasingly complex. The emergence of generative AI and the widespread circulation of misinformation, disinformation, and malinformation have created new challenges for public health authorities, researchers, and communicators alike.

With the ability to rapidly produce realistic text, images, and videos, generative AI has introduced the era of "synthetic realities"—defined as "any contextual digital creation or augmentation enabled by AI methods" (Cardenuto et al., 2023, p. 2). These AI techniques employ great amounts of data, leading to a new "reality" or narrative, regardless of its intention to deceive the individual or groups interacting with it. This poses significant risks to public trust, especially during health crises. Vulnerable populations—such as older adults, people with limited digital literacy, or communities with limited access to verified health sources—are particularly susceptible to the harmful effects of misleading information and synthetic creations. This development poses major challenges for governments worldwide. Therefore, in this article, we ask: How do governments in different countries address the challenges posed by AI? By analyzing Germany's and Brazil's approaches to combating information disorders through hard and soft law, the article aims to contribute to a deeper understanding of how public institutions can navigate the complexities of public health and emerging technologies. It is important to mention that, in this article, we understand information disorders as comprising three types: disinformation (false information deliberately created to harm a person, social group, organization, or country), misinformation (false information created without the intention to harm a person, social group, organization, or country), and malinformation (information based on reality but used to inflict harm on a person, social group, organization, or country; Wardle & Derakhshan, 2017).

In addition to the introduction and conclusion, the present article is divided into five parts. Sections 2 and 3 introduce the concepts of health risk communication and AI, forming the theoretical framework that underpins our analysis. Section 4 describes the methodological approach adopted and the corpus selection criteria, which are based on a document analysis of hard and soft law governance approaches. Finally, Sections 5 and 6 present and discuss the research findings, offering a comparative analysis of Brazil and Germany.

2. Health Risk Communication

Health risk communication can be defined as "the exchange of information among interested parties about the nature, magnitude, significance or control of risk" (Nicholson, 1999, p. 253). Considered a central aspect of human subjectivity in Western societies, risk can be understood as something "that can be managed through human intervention and is associated with notions of choice, responsibility, and blame" (Lupton, 1999, pp. 25).

The process of health risk communication is inherently bidirectional. Merely presenting information, without addressing the underlying complexities and associated risks, does not guarantee effectiveness. For risk communication to be successful, it is essential to inform the target audience about the existence of potential threats, assist them in understanding their vulnerabilities, and provide both guidance and support on strategies to reduce the likelihood of exposure and harm (World Health Organization, 2018). Ineffective communication in this domain not only fails to adequately prepare the public for such threats but may, in fact, exacerbate risks and potential harms by conveying inappropriate explanations and eliciting cognitive responses that act as barriers to effective action (European Monitoring Centre for Drugs and Drug Addiction, 2023).

The objectives of health risk communication are: "(1) to present information in such a way that it is understood and usable; (2) to ensure the audience is sufficiently informed so as to make sound judgments regarding risk; and (3) to engage the active support of those affected" (Nicholson, 1999, p. 253). Its critical domains of action include hazard and risk, the individual (perceptions), the social environment (group dynamics), and communication itself.

Health risk communication is shaped by a wide range of interrelated factors, as illustrated by the framework in Figure 1. These include individual conditions such as personal capabilities, values, and prior experiences; interactional dynamics, for instance in how information is exchanged between individuals or groups with different hierarchical standing, e. g., between physician and patients or more distanced communication formats; organizational factors, such as the structures and practices of health institutions, as well as the increased use of social media platforms; and finally, broader societal aspects, including cultural norms, media environments, and political systems. Together, these dimensions influence how health risk information is perceived, trusted, and acted upon (Ruhrmann & Guenther, 2025).

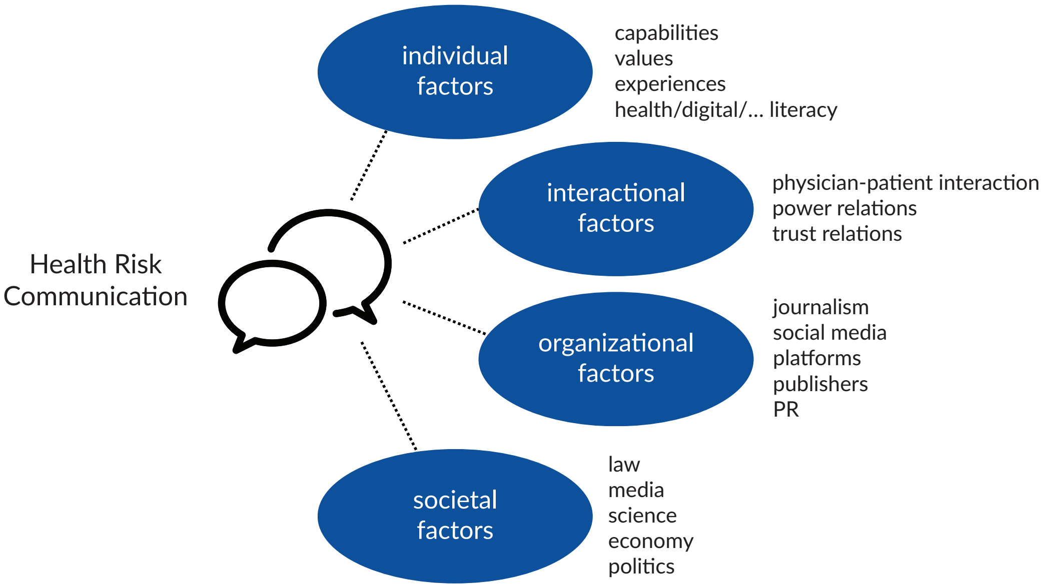

Figure 1. Health Risk Communication Framework. Source: Adapted from Ruhrmann and Guenther (2025).

2.1. Challenges

The rapid advancement of communication technologies has considerably impacted the consumption of information from mass media such as radio, television, and newspapers (Berg et al., 2021)—which were the main channels through which authorities communicated with the public. Online communication has transformed the way people seek information, enabling the entry of other actors into this ecosystem—actors who are not always committed to disseminating accurate and high-quality content (Monari, 2024). This has opened the way for the phenomenon of information disorders (Wardle & Derakhshan, 2017), which can undermine individuals' adherence to health information coming from governmental authorities or knowledge-producing institutions, such as science and journalism. A consequence of this is infodemics, which the World Health Organization defines as an overabundance of information—both accurate and inaccurate—that impairs the public's ability to identify reliable sources (World Health Organization, 2025).

This phenomenon, which can be exacerbated by the advancement of AI, is only one of the challenges faced by authorities in the field of health risk communication. This is because each type of audience responds differently to health threats and risk communication, revealing that multiple forms of health literacy exist in our society (Glik, 2007). Several factors can help explain these variations, such as: (a) imminence of potential harm (immediate consequences have a greater impact on behavior than long-term ones); (b) voluntariness of the action (voluntarily assumed risks are perceived as less severe); (c) perceived control over the risk (risks believed to be under an individual's control are perceived as less severe); (d) familiarity with the risk (familiar risks are perceived as less severe than new ones); (e) personal or social past experiences with the threat; and (f) how the risk is learned through modeling, filtering, and prioritization by communication channels, such as mass media (European Monitoring Centre for Drugs and Drug Addiction, 2023). Other factors can also shape risk perception, including gender, age, ethnicity, income, education, and health literacy levels. Therefore, the way people respond to health threats and harms depends not only on how they think about the issue, but also on how they feel about it and on who is addressing it with them (Peters et al., 2006). Health risk communication accordingly also encompasses societal and relational dimensions.

2.2. Vulnerable Groups

Vulnerable and minority groups are considered central audiences in health risk communication due to social, epidemiological, and communicational factors. Women, the elderly, adolescents, youth and children, people with disabilities, Indigenous peoples, refugees, immigrants, and minority groups are considered those with the highest degree of socioeconomic marginalization (UNICEF, 2020). This is because they are more exposed to risks (environmental, occupational, and social) as a result of structural inequalities and may face difficulties in understanding the messages disseminated by health authorities because of linguistic, cultural, or health literacy barriers (Winarnita et al., 2025). Additionally, some members of these groups carry a historical distrust toward institutions due to experiences of discrimination or negligence by authorities and/or health systems. Regarding the latter, Brazil presents an episode in its history that illustrates this fact: In 1904, the Vaccine Revolt took place, a widespread uprising triggered by public dissatisfaction with the mandatory smallpox vaccination campaign imposed by the federal government (Dandara, 2022). The revolt stemmed from the authorities' lack of information about the importance of vaccination.

Hence, health risk communication can play an important role in supporting vulnerable populations in their decision-making during health crises. It can help these groups, for instance, to increase their awareness of historical, social, and economic disadvantages (Kwek, 2017). Consequently, authorities must consider social implications when developing health risk communication strategies, such as limited access to resources, cognitive reasoning, and sociocultural factors, in order to design a more comprehensive plan for these populations (Eisenman et al., 2007).

An example of these implications lies in internet access, as online media are one of the communication channels used by institutions during health crises. It is understood that internet access is not a resource available to everyone, as it varies according to gender, age, location, and social class (Huxhold et al., 2020). In Germany, 94% of the population has internet access (World Bank, 2024); however, 2.8 million people (4% of the population) between 16 and 74 years old had never used this service by 2024, according to data from the German Federal Statistical Office (Bachels, 2025)—older adults (aged 65 to 74) represent the largest cohort in these statistics, which places them at risk of social exclusion during health crises, such as epidemics (Huxhold et al., 2020). In Brazil, 84% of the population has access to the internet (World Bank, 2024), but only 22% have satisfactory connectivity conditions (Núcleo de Informação e Coordenação do Ponto BR, 2024)—that is, they can effectively use various online services, allowing them to take advantage of opportunities in the digital environment. Internet access in Brazil is therefore unequal, with white men from higher social classes being those with the most satisfactory connectivity conditions (Núcleo de Informação e Coordenação do Ponto BR, 2024).

Moreover, it is necessary to consider that information alerts directed to vulnerable and minority groups should address aspects such as linguistic barriers and more limited options to receive warning messages (Burger & Gochfeld, 2019; Teo et al., 2018). It is essential to give them priority assistance and involve them in decision-making processes for response, recovery, preparedness, and risk and harm reduction.

Given the challenges of health risk communication, ranging from social and digital inequalities to information disorders, it becomes essential to consider the role of emerging technologies in this ecosystem. AI thus emerges as an ambivalent force: capable of personalizing messages and supporting decision-making, but also of reinforcing biases and amplifying the circulation of persuasive synthetic content. The following section will examine these potential applications and limitations in the context of health risk communication.

3. AI

AI refers to the capacity of computational systems to interpret data, learn from it, and apply that knowledge to achieve specific goals and tasks through flexible adaptation (Haenlein & Kaplan, 2019, p. 5). Data and algorithms form the foundation of contemporary AI. Data serve as the material from which AI systems derive patterns, correlations, and predictive insights, while algorithms provide the structured procedures that enable the processing, analysis, and transformation of this information into actionable outputs. In health communication, large and diverse datasets—ranging from epidemiological statistics to social media discourse—allow AI models to identify emerging public health concerns, segment target audiences, and tailor messaging strategies.

3.1. Chances

AI can be used in physician–patient interaction, as decision aid, personal health record interpretation tool, and tailored patient information resource (Antel et al., 2022; Miller et al., 2024). Further, AI can be used to efficiently create targeted and tailored health risk narratives customized to audience demographics and psychological factors (Chu & Liu, 2025; Miller et al., 2024), to real-time monitor public health indicators (e.g., vaccination rates), to automate routine tasks (e.g., content creation and dissemination), and to enhance real-time engagement through, for example, chatbots (Miller et al., 2024). Additionally, it serves as an important tool in emergency (health) communication and disaster management (Reynolds & Seeger, 2005).

The use of AI-powered language translation tools can facilitate communication between different languages, making health information more accessible to diverse populations (Miller et al., 2024).

3.2. Challenges

The integration of AI into health risk communication presents both technical and societal challenges. A central concern is the quality and representativeness of training data, as AI outputs are constrained by the completeness, accuracy, and diversity of the datasets they rely on. This links closely to the issue of algorithmic bias, whereby historical inequities embedded in data risk perpetuating or amplifying disparities in healthcare. Obermeyer et al. (2019), for example, found that a widely used health prediction algorithm in the US underestimated the needs of Black patients because it used past healthcare expenditure as a proxy for health status—a variable shaped by systemic underinvestment in minority health. This case illustrates the limits of correlation-based modelling: Statistical associations may misrepresent underlying realities when social determinants remain unaccounted for.

Another relevant issue is the alignment of AI with the cultural and ethnic aspects of each society. Ogie et al. (2018) reinforce the need for research that considers the development of “culturally intelligent AI machines,” that is, machines capable of generating disaster risk communication (e.g., emergency alerts) based on learned culture and communication protocols practiced in different linguistic and ethnic groups within society.

Beyond technical concerns, AI-driven health risk communication must contend with a media landscape increasingly distorted by infodemics. Generative AI models significantly amplify the potential for mis-, dis-, and malinformation by enabling automated and rapid production of false content at unprecedented scale (Cardenuto et al., 2023). Large language model (LLM) systems can generate persuasive, well-written narratives, disseminating misleading information and creating a new reality: “a synthetic reality” (Cardenuto et al., 2023). This automation increases both the speed and volume of information disorders circulating online, including in highly sensitive areas such as medicine, where inaccurate health claims can have severe consequences. The sophistication of AI-generated information disorders heightens their potential to appear credible, thereby increasing the likelihood of public acceptance and further dissemination, particularly on social media platforms—especially since this type of content exploits cognitive biases “(e.g., anchoring bias, third-person effect, authority bias, bandwagon effect, to mention a few), which are systematic errors in judgment that humans can make” (Cardenuto et al., 2023).

Given the theoretical premises presented, the next section will demonstrate the methodological perspective of this article, which is structured around a document analysis (Moreira, 2017) based on the READ approach (Dalglish et al., 2020).

4. Methods

Document analysis can be defined as the process of identifying, verifying, and assessing documents for a specific purpose (Moreira, 2017). It involves a series of structured steps. In this study, we employed the READ approach (Dalglish et al., 2020), which consists of four stages: (a) read your materials, (b) extract data, (c) analyze data, and (d) distill your findings. According to Dalglish et al. (2020), in the first step researchers should establish research parameters—such as the nature and number of documents to be analyzed—based on the research question. Our research period was between March and July 2025, and we initially chose to analyze 15 documents (see Supplementary File 1). They were selected based on their demonstrated governmental impact and institutional relevance. These sources were chosen because they represent key policy instruments and official communications that have directly influenced the development and implementation of AI health regulation in Germany and Brazil. The selection aimed to capture a balanced representation of perspectives from different governmental and institutional levels (hard law and soft law).

Academic publications and internal company regulations were deliberately excluded in our selection in order to maintain the focus on governmental and institutional documents. Including them would have gone beyond the intended scope and objectives of the analysis. Consequently, the final selection prioritizes documents that are both authoritative and directly connected to decision-making structures, ensuring the analysis remains grounded in materials with verifiable institutional significance.

The second step is data extraction, which involves organizing the collected documents so that the information can be read coherently and data specifically related to the research question can be extracted. In this phase, the previously established Health Risk Communication Framework (see Figure 1) served as the analytical foundation.

The third stage, data analysis, builds directly on the previous one. Once the documents have been organized and read, researchers begin to form preliminary theories that may later be confirmed or refuted as they complete this stage and gain a "full picture." In our study, we adopted a case study methodology combined with a comparative analysis based on hard and soft law perspectives in two different countries (see Supplementary File 2).

Our decision to conduct a comparative study between Brazil and Germany was guided by the fact that the use of AI across various sectors of society—including health risk communication—is a global issue. Countries have adopted different approaches to address it, whether through guidelines, implementation plans, or legislation. Both Brazil and Germany are democracies; however, while Germany follows a corporatist model, Brazil is characterized by liberalism (Cazzamatta & Sarısakaloğlu, 2025). These differing regime types shape how state and societal actors interact in policy-making and, by extension, how emerging technologies such as AI are governed. In corporatist democracies, governance tends to be consensus-oriented and involves organized stakeholder participation (e.g., medical associations and professional bodies), which may result in negotiated, more deliberative, and potentially slower adoption of AI tools. Liberal democracies, in contrast, often emphasize market-driven innovation and individual rights, which can enable faster but more fragmented governance responses. As a result, they represent distinct trajectories of technological governance shaped by different sociopolitical contexts. These contrasts make them particularly compelling cases for comparison. The choice was also motivated by the authors' nationalities, as one is Brazilian and the other German.

The final stage of the READ approach (Dalglish et al., 2020), distilling your findings, marks the conclusion of the document analysis. In our case, reaching saturation enabled us to synthesize the material and provide a coherent overview of the most significant hard and soft law regulations shaping AI health risk communication. The conceptual framework outlined in Sections 2 and 3, namely the definition of health risk communication and AI, ultimately forms the basis for the document analysis.

5. Findings: Governance Approaches to Address Challenges Raised by AI in Health Risk Communication

The analysis is structured along two main dimensions to ensure comparability between the cases: hard law and soft law (European Center for Constitutional and Human Rights, 2025). Information disorders are addressed through different but interrelated governance strategies. National governments rely on hard law or statutory regulation, often embedded within multi-level governance structures involving international institutions such as the European Union (EU). These legal measures are complemented by alternative modes of governance referred to as *soft law*, which include the public health organizations, fact-checking organizations, collective self-regulation within industry sectors (Saurwein & Spencer-Smith, 2020), and various forms of civil society initiatives.

In addition to this regulatory distinction, the analysis draws on the Health Risk Communication Framework (see Figure 1), which provides the analytical foundation for examining how hard and soft law instruments influence the communication of health risks, particularly in the context of AI and information governance.

5.1. Germany

In the EU and in Germany, regulatory governance of digitally amplified information disorders has focused predominantly on social media platforms, which are framed as the primary vectors for the circulation of harmful or misleading content (European Commission, 2018; Saurwein & Spencer-Smith, 2020). Germany’s legislative framework for regulating digital platforms and online communication increasingly draws upon the provisions of the EU’s Digital Services Act (DSA), adopted in 2022.

The DSA is a landmark EU regulation enacted in 2022 to modernize and supersede the earlier e-Commerce Directive (Frosio, 2024). It establishes unified obligations across all EU member states for digital platforms—ranging from online marketplaces to social media networks and hosting services—that act as intermediaries for disseminating content within the EU. The primary objectives of the DSA are twofold: to create a safer and more transparent digital environment that upholds users’ fundamental rights, and to cultivate a fair and competitive digital single market. Specifically, platforms—particularly very large ones (namely AliExpress International, Amazon, Meta, etc.; cf. European Commission, 2025a)—must comply with transparency requirements, disclose how their algorithms and content moderation systems function, create user-accessible mechanisms for content reporting and complaints, inform users about chances to the terms of service, and manage systemic risks such as the spread of mis/disinformation. In short, the DSA seeks to rebalance the digital landscape by imposing legally binding standards that ensure greater accountability and safety in online environments. The document is a measure of hard law that is legally binding for all EU member states from 2024.

Since July 2025, the European Commission has incorporated the voluntary Code of Conduct on Disinformation, first launched as Code of Practice on Disinformation in 2018 and updated in 2022, into the DSA (Jahangir, 2025a). The Code now serves as a key compliance benchmark for online platforms, shifting from a non-binding framework to a potential basis for enforcement. It addresses six areas: (a) demonetization, by avoiding advertising placements next to disinformation and improving cooperation among actors; (b) transparent political advertising, through clearer labelling and stronger transparency obligations; (c) reduction of manipulative behavior, targeting both existing and emerging forms of manipulation while fostering cooperation among signatories; (d) user empowerment, by developing better tools and facilitating access to reliable information and context; (e) fact-checking coverage throughout the EU, ensuring consistent use of fact-checkers' work and providing fair financial support; and (f) data access for research, granting researchers easier access to platform data and backing scientific inquiry. Taken together, these provisions reflect the EU's attempt to create a comprehensive, multi-level framework to mitigate disinformation (see Figure 2). The Code of Conduct was signed by major online platforms and search engines such as Google, as well as smaller online platforms, the advertising industry, fact-checkers, civil society, research organizations, and players offering technical solutions.

However, the effectiveness of the DSA is deeply embedded in the corporatist political structures of the European system, which serve as its precondition. As a result, some of its solutions are viewed as *too European* in scope and design and may not be readily transferable to non-European contexts (Husovec, 2023). Nevertheless, some scholars expect that the high-level principles embedded in the DSA could provide a foundation for dialogue among liberal democracies on how best to regulate user-generated content services (Husovec, 2023).

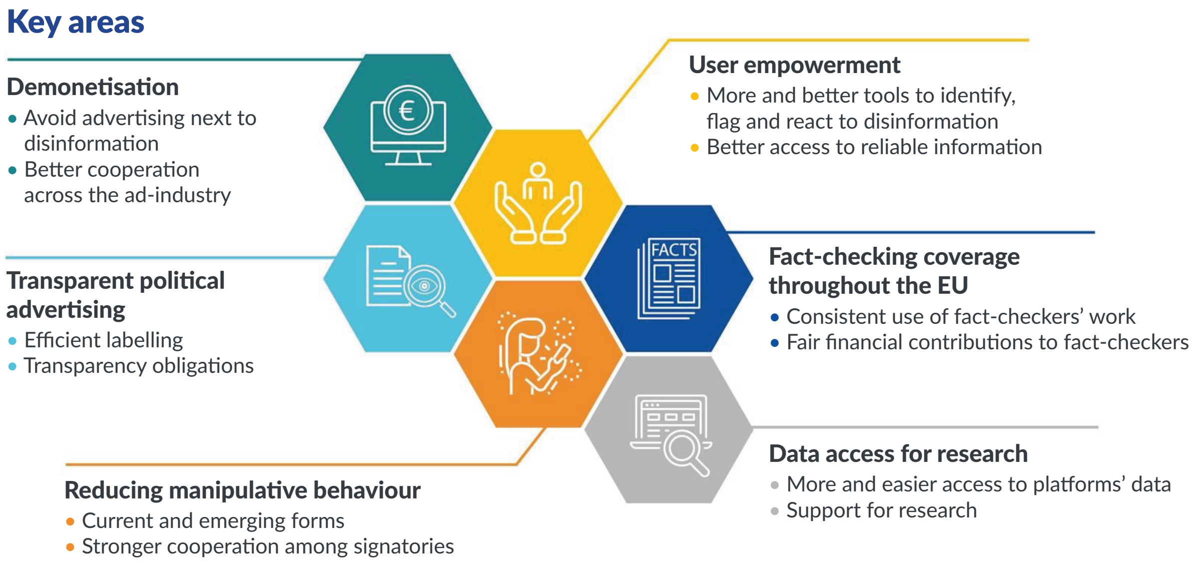

Figure 2. Key areas of the Code of Conduct on Disinformation. Source: European Commission (2025b).

The DSA forms the basis of the German national law Digitale-Dienste-Gesetz (DDG). Compliance with the DDG is monitored by the Federal Network Agency as Digital Service Coordinator (DSC; Bundesnetzagentur, 2025). The DSC receives reports of possible DSA violations primarily through complaints from users. If such reports, usually supplemented by further investigations, are sufficiently substantiated and relate to service providers based or legally represented in Germany, the DSC initiates proceedings. By the end of the 2024 reporting period, the DSC had initiated a total of four administrative proceedings against service providers (Bundesnetzagentur, 2025). Three of the four proceedings concern possible violations of DSA requirements for the establishment of reporting and redress procedures (Art. 16 DSA), the justification of measures against users (Art. 17 DSA), and the design of the platforms' internal complaint management system (Art. 20 DSA; Bundesnetzagentur, 2025). The DSC actively engages with the newly established European Board for Digital Services, which brings together the European Commission and all national DSCs (Bundesnetzagentur, 2025).

The DDG explicitly mentions that the provision of digital services (excluding audiovisual media services) by a provider established in another EU member state may be restricted under German law if such measures are necessary to protect public health, among certain other fundamental interests (Section 3 DDG).

However, in January 2025, Meta announced it would discontinue its professional, third-party fact-checking program in the US and instead adopt a user-driven "community notes" model—akin to the system used by Elon Musk's platform X (Graves, 2025). Community Notes is a user-driven fact-checking system in which participants contribute explanatory notes to flag and contextualize potentially misleading posts on X/Twitter. Yet, empirical evidence casts doubt on the effectiveness of this approach. Especially for tweets from verified users with many followers, it has not significantly reduced engagement with misinformation in terms of retweets or likes (Chuai et al., 2024). Instead, findings suggest that crowdsourced fact-checks often arrive too late to curb the viral spread of misleading content in its most influential early stages (Chuai et al., 2024). The change has raised alarm among fact-checkers, who viewed Meta's decision as politically motivated and a serious threat to informed public discourse (Graves, 2025). Further, Meta's undertaking to replace independent and professional fact-checking with community notes would violate the DSA if applied to the EU (Zingaretti et al., 2025). However, scholars are not very optimistic that the DSA is applicable to large companies such as Meta since charges brought against X remain unresolved (European Commission, 2023; Jahangir, 2025b; Kahn, 2025). The case illustrates that national and international legal frameworks, still in a formative stage, are currently undergoing a significant stress test.

Next to the hard law, there are various stakeholders who propose several soft law measures to combat information disorders. For the purpose of this study, we focused specifically on national public health authorities, independent fact-checking organizations, and NGOs or public initiatives, as these actors play a central role in shaping non-binding yet influential frameworks for communication governance in the health domain.

The Robert Koch Institute is one of the most important institutions for health risk communication in Germany. As the national public health institute, it plays a central role in monitoring, communicating, and managing health-related risks, particularly during crises such as pandemics. The Robert Koch Institute regularly publishes reports, commission statements, and scientific articles on its website and in academic journals, contributing substantially to evidence-based public health communication. However, regarding the intersection of AI and health risk communication, we identified one publication that explicitly draws on Robert Koch Institute data to explore ways of improving infodemic management during health crises. The study by Boender et al. (2023) proposes that one effective way to enable people to make informed health decisions is through responsive, evidence-based, and audience-specific risk and health communication. The authors state that well-designed and carefully implemented infodemic management can play a key role in tailoring the right messages to the right audiences at the right time, while simultaneously strengthening both health literacy and scientific literacy (Boender et al., 2023). In their opinion, this includes "social listening," which could be enhanced by AI, for example by tracking health discussions on social media (Boender et al., 2023). Social listening is the process of systematically extracting and analyzing information from social media channels to monitor social developments in real time. It involves detecting emotions, topics, and opinions; mapping patterns of information flow; and modeling opinion networks to understand how ideas and sentiments spread across digital spaces (McGowan, 2022; Purnat et al., 2022).

Further, independent fact-checking organizations play an important role in addressing misinformation alongside regulatory initiatives in Germany. Prominent examples include Correctiv, a non-profit investigative newsroom that operates a fact-checking unit, and the Deutsche Presse-Agentur, which runs its own fact-checking service as part of its news operations. Both organizations monitor and verify claims circulating in public discourse, particularly on social media, and publish corrections to counter false or misleading content. Their work illustrates how non-state actors complement formal governance frameworks by providing expertise, credibility, and public-facing resources in the fight against information disorders. The German fact-checking organization Correctiv, as well as the Deutsche Presse-Agentur, both adhere to the International Fact-Checking Network (IFCN) Code of Principles (IFCN, 2025). This Code establishes a set of professional and ethical commitments for organizations that regularly publish non-partisan reports assessing the accuracy of statements made by public figures, institutions, and other widely circulated claims related to matters of public interest. Developed through global consultations among fact-checkers, the Code emphasizes transparency in sources, funding, methodology, and corrections, as well as a commitment to impartiality and fairness. Correctiv's adherence to the IFCN Code underscores its role as an independent and credible actor in the landscape of information governance and infodemic management in Germany. Its activities contribute to maintaining trust in public communication and serve as a soft-law mechanism promoting accountability and accuracy in the dissemination of health-related information.

Additionally, German public institutions try to involve a broad public in developing governance guidelines. For example, the initiative of the Bertelsmann Stiftung "Forum gegen Fakes—Together for a Strong Democracy" (Bertelsmann Stiftung, 2025a) addressed the general public and offered multiple opportunities for active participation between January and June 2024. In three online participation phases, citizens were invited to contribute their perspectives online, propose measures to counter disinformation, and vote on ideas submitted by others. At the same time, a citizens' assembly of more than 120 randomly selected and socio-demographically diverse participants worked on evaluating and refining these proposals. This resulted in 15 policy recommendations comprising 28 concrete measures (Bertelsmann Stiftung, 2025a). The final proposals submitted clearly indicate where participants' priorities lie in addressing information disorders. Raising awareness and strengthening education were considered particularly important, for example through the promotion of media literacy, information campaigns, and communication tailored to specific target groups. Participants also emphasized the need for changes in media practices, journalism, and the functioning of social networks. For example, a clear majority (73%) agreed that content produced by AI should always be explicitly labeled as such. In addition, participants supported the idea of introducing a certification scheme for media outlets that provide additional source information about their content, underscoring the value placed on credibility and accountability in news and information practices. There was broad consensus that key actors such as policymakers, media organizations, and platform providers must assume greater responsibility, while at the same time individual responsibility was seen as a crucial element in dealing with disinformation. Some participants also called for more restrictive measures, such as political and regulatory interventions, whereas others regarded such steps as unwarranted infringements on the freedom of speech (Bertelsmann Stiftung, 2025a).

5.2. Brazil

The Federal Constitution of Brazil (Constituição da República Federativa do Brasil, 1988) establishes health as a universal right and a duty of the State. Law No. 8,080 of 1990, commonly referred to as the "SUS Law," defines the conditions for the promotion, protection, and restoration of health and formally institutes the Unified Health System (Sistema Único de Saúde, SUS), a public framework designed to guarantee universal access to healthcare for all Brazilian citizens (Presidência da República do Brasil, 1990). Both the Federal Constitution and the SUS Law recognize access to information as a fundamental principle, assigning to SUS the responsibility of disseminating information, particularly in contexts of public health risks.

Although informing, guiding, and engaging society in times of health risks—such as epidemics, environmental disasters, and health emergencies—are already provided for in Brazilian law, the country still faces challenges in the field of health risk communication. Regarding the hard law, these include combating mis/disinformation, ensuring accessibility, addressing the dilemma between technical communication (based on scientific evidence) and the political use of information (narrative shifts depending on governments), and the effective implementation of the principle of information transparency. The emergence and integration of AI may either exacerbate or help mitigate these challenges, depending on how it is regulated and applied.

The Brazilian government has not yet established specific legislation or regulations for AI environments. Law No. 2,338 of 2023, already approved by the Federal Senate, addresses the use of AI in the country but is currently under review by the House of Representatives (Senado Federal do Brasil, 2023). This legislative proposal is oriented toward the development, implementation, and responsible use of AI systems, grounded in the principles of prevention, precaution, and the mitigation of systemic risks derived from both intentional and unintentional uses, as well as from unforeseen effects of such systems. Among its provisions, the bill foresees the creation of a competent authority to oversee compliance with best practices in AI governance; however, it does not specify which institution would be responsible for such oversight—whether a newly established entity or an existing one.

This proposed law exhibits several characteristics typical of the liberal democracy model, as it establishes individual rights, transparency, state oversight, equality, and free enterprise as the foundations of AI governance in Brazil. Although not explicitly mentioned, it also supports health risk communication by emphasizing transparency, auditability, and protection against bias. However, it does not mandate the creation or implementation of measures ensuring accessible language, public engagement, or emergency protocols—essential, for instance, to effectively and reliably inform the population during health risk situations such as epidemics and pandemics. Thus, while it prioritizes individual needs, it does not account for the environment in which individuals live, which may involve social and economic vulnerabilities. In a country as large, diverse, and unequal as Brazil, this is an essential consideration to make.

It is also important to mention that Brazil currently lacks comprehensive regulation of digital platforms. The Civil Rights Framework for the Internet (Law No. 12,965 of 2014) constitutes the primary legal framework establishing principles, guarantees, rights, and obligations for internet use in the country (Presidência da República do Brasil, 2014). It safeguards freedom of expression, user privacy, and network neutrality, while also defining the role of the State in internet-related matters. Nevertheless, given that it was enacted more than a decade ago, many scholars and policy experts argue that this framework is insufficient for addressing contemporary challenges such as mis/disinformation and hate speech, as well as inadequately considering technological advances brought by AI. An illustrative example of this is Article 19 of the framework, which conditioned the civil liability of digital platforms on the existence of a prior judicial order for content removal.

This article was the subject of judicial review by the Federal Supreme Court (Supremo Tribunal Federal, STF), Brazil's highest judicial entity. In June 2025, the STF ruled on the partial constitutionality of Article 19, determining that platforms and internet providers could henceforth be held liable without a prior judicial order in two specific contexts: (a) in cases of boosted or paid content, including dissemination by bots or artificial distribution networks; and (b) in instances involving the massive circulation of highly harmful content, such as anti-democratic acts, terrorism, incitement to suicide or self-harm, racial, religious, sexual, gender-based, or misogynistic hate speech, child pornography, sexual crimes against vulnerable individuals, and human trafficking ("STF define parâmetros," 2025).

In terms of national initiatives in the field of AI, the federal government launched the Brazilian Artificial Intelligence Plan (Plano Brasileiro de Inteligência Artificial, PBIA) in 2024 (Conselho Nacional de Ciência e Tecnologia, 2024). The plan's objective is described as follows:

> Promote the development, availability, and use of AI in Brazil, oriented toward addressing major national, social, economic, environmental, and cultural challenges, in order to safeguard individual and collective rights and security, foster social inclusion, defend democracy, ensure information integrity, protect labor and workers, safeguard national sovereignty, and promote the sustainable economic development of the nation. (Conselho Nacional de Ciência e Tecnologia, 2024, p. 20)

The PBIA distinguishes between short-term actions (initiatives already underway or to be implemented soon) and long-term structural actions. While it does not explicitly address health risk communication, two initiatives outlined in the plan may be interpreted as relevant to this domain. The first, categorized as a short-term action, is the deployment of Generative AI for Personalized Health Care. This AI assistant is designed to optimize the personalization of health care within the framework of digital primary health services and will be integrated into the SUS. The second, a structural measure, is the AI Dissemination and Outreach Program, intended to foster digital literacy and public engagement with AI by popularizing technological foundations, transparency, everyday applications, risks, and citizens' rights. To achieve these goals, the government plans to establish a Brazilian AI Olympiad, develop partnerships with universities and technology firms for workshops and mentoring, and promote public information and educational campaigns on AI.

In the field of health risk communication, the Brazilian Ministry of Health has implemented a noteworthy initiative: the Health with Science (Saúde com Ciência) program. It seeks to promote and strengthen public health policies and science through actions to identify and understand the phenomenon of information disorders and provide society with accurate information. One of the project's actions is a fact-checking service, where people can send content they suspect to be false for verification (Brazilian Ministry of Health, 2025). Although this initiative may be well-intentioned, it raises concerns about conflicts of interest and credibility. This stems from the fact that communication originates from a biased perspective—that of the government, the political party currently ruling the country—and is therefore shaped by the interests of those in power (Meseret, 2024).

Similarly to Germany, Brazil also has different actors proposing soft law measures to address the challenges of health risk communication and to combat information disorders. One of them is Fiocruz (Oswaldo Cruz Foundation), a health research institute affiliated with the Ministry of Health and the SUS, which offers various initiatives as well as high-quality information during times of crisis, such as the Covid-19 pandemic.

Among the recent guidelines and strategies adopted by the institution are: the promotion of dialogues and training programs; cultural and science communication initiatives; scientific research and analyses of social representations; and the production of guides and recommendations. Regarding the latter topic, Fiocruz launched a guide to combat health misinformation aimed at health professionals in 2024, in partnership with other institutions. The document provides tips on how to deal with this situation in their daily professional lives (Sacramento et al., 2024).

In addition, Brazil has a number of well-established independent, non-governmental fact-checking services, such as Lupa and Aos Fatos (Lelo, 2022), which provide systematic verification of information and contribute to raising public awareness. These initiatives illustrate how non-state actors in both countries complement formal governance frameworks by providing expertise, credibility, and public-facing resources in the fight against mis/disinformation (Cazzamatta & Sarısakaloğlu, 2025). Both Lupa and Aos Fatos have signed the IFCN Code of Principles and are subject to regular audits by the organization to ensure they continue to adhere to its principles, such as impartiality and transparency (IFCN, 2025).

The country has also a virtual collective organization called the National Network to Combat Disinformation (Rede Nacional de Combate à Desinformação), which connects projects and institutions of different kinds that work and contribute in some way to combating the disinformation market—such as communication collectives, universities, educational communication projects, fact-checking projects, research projects, etc. (Rede Nacional de Combate à Desinformação, 2025). Among the members of the Network is the Democracy in Check Institute (Instituto Democracia em Xeque), whose mission is to produce knowledge to combat disinformation campaigns, hate speech, and violent political extremism (Instituto Democracia em Xeque, 2025).

Although Brazil has implemented effective health risk communication measures, both through hard and soft law mechanisms, the country still struggles with the politicization of information and the lack of accountability of political actors and social media platforms for the dissemination of false or misleading content. During the Covid-19 pandemic, the then president Jair Bolsonaro downplayed the virus's impact in favor of an economic agenda opposing social isolation and lockdown measures, promoted the use of unproven drugs against the disease, and spoke out against mandatory vaccination—going so far as to claim that those who received the vaccine could "turn into alligators" (Boschiero et al., 2021; Ricard & Medeiros, 2020). Also, in an attempt to undermine the media's efforts to inform the public about the spread of Covid-19 in the country, his government began releasing case and death figures with delays, making it impossible for news outlets to publish timely data. In response, journalists and news agencies in Brazil organized a consortium to provide reliable data and information about the virus, its societal impact, and recommended treatments (Bertelsmann Stiftung, 2025b). The Brazilian case demonstrates a government that not only neglects transparency-based measures but also instrumentalizes disinformation and hate campaigns as deliberate tools of political communication (Ozawa et al., 2023)

6. Comparative Discussion

6.1. Organizational Regulations and Individual Responsibility as Core Governance Pillars

Five observations were made when analyzing governance approaches to health risk communication in the age of AI in Brazil and Germany (see Supplementary File 2).

First, Brazil and Germany face similar challenges in the field of AI and health risk communication, particularly regarding the management of information disorders. Information disorders arise from socio-technical assemblages, consisting of platforms, actors, and big data, as well as political and commercial incentives (Saurwein & Spencer-Smith, 2020).

Second, governance approaches reveal different emphases and directions. Central to both countries is the mitigation of mis-, dis-, and malinformation (Bertelsmann Stiftung, 2025b, p. 6), which remains a core concern across regulatory initiatives. In addition, governance efforts focus on the regulation of digital intermediaries, the promotion of transparency, and the strengthening of media pluralism; within the European context, particularly through the DSA (Bertelsmann Stiftung, 2025b, p. 6).

Third, in both countries' governance approaches, AI is consistently portrayed as an ambivalent tool in health risk communication. On the one hand, it is identified as a driver of information disorders, enabling the large-scale production, personalization, and dissemination of misleading or false content. On the other hand, governance frameworks also emphasize AI's potential as a solution, highlighting its use in detecting, filtering, and countering such content. This dual framing positions AI simultaneously as a risk factor and a key instrument in efforts to mitigate the very challenges it creates—the PBIA illustrates this by positioning AI as a way to help solve various health problems faced by the SUS. In both countries, regulation frames AI governance as a socio-technical assemblage that must be addressed across multiple regulatory levels (Waisbord, 2020).

Fourth, with regard to the Health Risk Communication Framework, our analysis revealed that both the German and Brazilian governance approaches emphasize organizational and individual factors of health risk communication (see Supplementary File 1). However, the analysis shows that organizational regulations repeatedly reach their limits—either where they cannot be enforced and thus not upheld in court, or where their enforcement (for example, through established fact-checking mechanisms) is not sufficiently persuasive to effectively curb the spread of misinformation (Waisbord, 2020).

For example, in Brazil, the debate on the regulation of social media faces political and economic obstacles. This is because the platforms have acted vigorously to prevent bills on the matter from advancing—for instance, Google spent more than half a million reais on advertisements opposing the so-called Fake News Bill (Brazilian Law on Internet Freedom, Responsibility, and Transparency), seeking to influence public opinion and promote a negative view of the proposal (Fonseca, 2023; Pacheco, 2023). Far-right politicians are also opposed to regulation and align themselves with big tech interests. Arguing that "Brazil is under censorship" and that decisions made by the STF (such as the ruling on the Civil Rights Framework for the Internet) undermine freedom of speech in Brazil and the US, supporters of the former president Bolsonaro—and Bolsonaro himself—came out in defense of the trade tariffs imposed by US President Donald Trump on the country (Menezes & Mangabeira, 2025). A comparable tension can be observed in the EU, where the governance of health risk communication is also shaped by tensions surrounding freedom of expression and the power of digital platforms. While instruments such as the DSA establish new obligations for very large platforms, questions remain about how these measures can be effectively enforced in practice. Major technology companies wield significant legal and economic power, making it difficult to hold them accountable in court. At the same time, debates over free speech complicate regulatory efforts, as measures designed to curb harmful content risk being perceived as undue restrictions on fundamental rights. This ongoing struggle illustrates the delicate balance between safeguarding democratic discourse and ensuring that platforms do not become unregulated spaces for the spread of misinformation and disinformation.

Fifth, in both countries, governance approaches do not rely solely on regulating intermediaries but also emphasize strengthening individual competencies and personal responsibility in dealing with information (see Supplementary File 1). The aim is to enable citizens to recognize disinformation, engage in critical reflection, and make informed decisions. In Germany, this is primarily reflected in initiatives on media education, transparency, and the promotion of digital health literacy. In Brazil, the emphasis on individual responsibility is even more pronounced (Tabak et al., 2025). This can likely be attributed to the political context and comparatively weaker regulation of digital spaces, which has led to a stronger reliance on educational initiatives and the cultivation of individual resilience against misinformation

6.2. Governance Blind Spots: Interactional Dynamics and Societal Contexts

While governance approaches in both countries place particular emphasis on organizational and individual dimensions of health risk communication, they fail to sufficiently address several other critical components of health risk communication. In particular, vulnerability-related, societal, and interactional factors receive far less attention, despite their central role in shaping how health risks are communicated, perceived, and acted upon. By overlooking these factors, governance strategies risk narrowing the scope of health risk communication to individual responsibility and institutional action, without adequately accounting for the social dynamics and intersectional vulnerabilities that shape the reception and impact of health information.

First, addressing individual vulnerabilities requires acknowledging that audiences are not homogeneous but situated within diverse individual, structural, algorithmic, and interactional circumstances. This includes variations in education, health status, and access to information or technological infrastructure, all of which shape people's capacity to understand and respond to health risks. Moreover, individuals draw on their own values and lived experiences when interpreting health information, influencing whom they trust, how they assess credibility, and which risks they perceive as relevant. In addition, individuals interact with social media platforms in different ways, navigating environments where content is increasingly curated, ranked, and personalized through automated systems. Such dynamics can amplify existing inequalities by exposing some groups to higher levels of misleading or harmful information (Ott, 2025). For this reason, emerging technologies—and their impact on health risk communication—cannot be understood in isolation from the socio-economic and institutional contexts in which they operate (Cazzamatta & Sarısakaloğlu, 2025).

Second, in both countries, risk governance strategies disregard the symbolic meanings and the ways in which social and cultural contexts shape the understanding and negotiation of risk. From a Foucauldian perspective, Lupton (1999) argues that risk functions as a governmental strategy of power regulation, whereby populations and individuals are monitored and directed in accordance with the objectives of neoliberalism. Thus, risk becomes a moral technology, and the discourse surrounding it operates in alignment with the regulation of the body, contributing to the constitution of the "self" and of subjectivity in contemporary society. Within neoliberal thought, responsibility for health is increasingly individualized, framing well-being as a matter of personal care and self-management. This perspective, however, neglects the crucial dynamics of power, trust, and social relations that shape health outcomes.

It is important to mention that the crisis of confidence in epistemic institutions (Monari, 2024) is part of the post-truth scenario and is something used by agents of disinformation to spread false information. For Waisbord (2018), the post-truth scenario has revealed the capacity of the digital world to present different perspectives on the same issue: "Social media platforms and search companies provide plenty of space for epistemologies with varying relations with reality. More interested in traffic and profit than in channeling scientific ideas, they offer opportunities to disseminate content disconnected from scientific truth-telling" (Waisbord, 2018, pp. 5–6). This undermines public health campaigns, especially when they are incorporated into political discourse—as the Brazilian example demonstrated. In this context, holding political actors accountable could be a way to combat information disorders and is something that could be done through international efforts, international organizations, or supranational bodies.

Third, our analysis indicates that effectively addressing information disorders in health risk communication requires the integration of interactional factors—ranging from changes in mutual relationships (e.g., between information sources and audiences, physician–patient interaction, etc.) to international interaction scenarios. We must understand health risk communication in the age of AI as a socio-technical assemblage. This means recognizing that communication processes are not shaped solely by individual behaviors or organizational strategies, but by the complex interaction between human actors, technological systems, institutional structures, and wider social dynamics. Saurwein and Spencer-Smith (2020) identify the following actors as being involved in information disorders on social media: (a) social media companies, who are responsible for platform design and policies; (b) publishers of mis-, dis-, or malinformation; (c) users who distribute the information by clicking, watching, and sharing; and (d) non-human actants, such as AI, bots, and algorithms.

Further, the global nature of digital platforms enables false or misleading health content to cross borders almost instantaneously, diminishing the effectiveness of isolated, nation-specific interventions. Misinformation originating in one country can rapidly influence public perceptions and behaviors elsewhere, especially when shared languages, cultural connections, or diaspora networks facilitate its spread. Given these dynamics, effective countermeasures demand coordinated international actions—such as harmonized regulatory standards, cross-border fact-checking collaborations, and joint digital literacy initiatives—to address the inherently global and decentralized nature of the problem. In healthcare, this is even more critical, as misinformation can lead people to make decisions that put their own lives and others' lives at risk, such as refusing vaccinations, adopting treatments and medications without scientific evidence, delaying diagnoses, or even abandoning safe therapies.

7. Conclusion

In summary, this comparative analysis of Brazil and Germany demonstrates that, while the governance of health risk communication in the age of AI operates through distinct institutional logics, both countries grapple with similar structural challenges. The interplay between hard and soft law mechanisms reflects an ongoing tension between regulation, technological innovation, and freedom of expression. Across both contexts, AI emerges as a double-edged instrument—simultaneously amplifying the risks of misinformation and offering new opportunities for its detection and mitigation. Yet current governance frameworks remain largely focused on organizational and individual dimensions. Within these frameworks, the individual dimension is primarily interpreted in terms of responsibility. As a result, individual vulnerabilities and interactional factors that shape trust and communication practices are often overlooked.

To move beyond this narrow focus, future governance approaches should adopt a more holistic and globally coordinated perspective, one that integrates interactional dimensions of communication. In this regard, communication and public trust emerge as central pillars of AI governance in health contexts, shaping how risks are perceived, negotiated, and acted upon. Building trustworthy systems requires more than technical regulation—it demands transparent communication practices, accountability mechanisms, and sustained dialogue between institutions, experts, and the public. Further, strengthening cross-border collaboration, harmonizing regulatory standards, and embedding social and ethical reflection into AI governance will be crucial steps toward building resilient and trustworthy health information ecosystems. Only through such an integrative approach can societies effectively address the evolving challenges of information disorders and ensure that technological progress contributes to, rather than undermines, public health.

Finally, we acknowledge that, although document analysis is appropriate for the objectives proposed in this study, it does not allow for inferences regarding public reception of the initiatives to combat information disorders in both countries. Interviews with governmental agents, non-governmental actors, fact-checkers, and the general public could be explored in future research to better understand how they assess the role of AI in exacerbating or mitigating this phenomenon in the health domain, as well as their perceptions of the use of AI in scenarios involving population-level risks. Accordingly, we understand that such future studies could broaden the empirical and temporal scope of our analysis, deepening knowledge on the subject and its evolution over time.

Acknowledgments

We would like to thank Deborah Lupton for inviting both of us as visiting scholars to her institute at the UNSW (Sydney, Australia) in 2023. By introducing us to each other, she made this article possible. We are also grateful to Fiona Bendig for her careful proofreading of the manuscript.

Funding

Publication of this article in open access was made possible through the institutional membership agreement between the University of Vienna and Cogitatio Press.

Conflict of Interests

The authors declare no conflict of interests.

Data Availability

The data analyzed in this study are listed in the Supplementary Files and are publicly accessible.

LLMs Disclosure

Generative AI (ChatGPT, GPT-4, OpenAI) was used to assist with language editing and stylistic refinement of the manuscript; the authors retained full control over the content and are responsible for all aspects.

Supplementary Material

Supplementary material for this article is available online in the format provided by the authors (unedited).

References

Antel, R., Abbasgholizadeh-Rahimi, S., Guadagno, E., Harley, J. M., & Poenaru, D. (2022). The use of artificial intelligence and virtual reality in doctor-patient risk communication: A scoping review. *Patient Education and Counseling*, *105*(10), 3038–3050. https://doi.org/10.1016/j.pec.2022.06.006

Bachels, C. (2025, April 11). Nearly 3 million Germans have never used the internet. *DW*. https://www.dw.com/en/nearly-3-million-germans-have-never-used-the-internet/a-72196617

Berg, S. H., O'Hara, J. K., Shortt, M. T, Thune, H., Brønnick, K. K., Lungu, D. A., Røislien, J., & Wiig, S. (2021). Health authorities' health risk communication with the public during pandemics: A rapid scoping review. *BMC Public Health*, *21*(1), Article 1401. https://doi.org/10.1186/s12889-021-11468-3

Bertelsmann Stiftung. (2025a). *Forum gegen Fakes—Gemeinsam für eine starke Demokratie. Ergebnisse der Evaluation des Bürgerrats und der Online-Beteiligung*. https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/forum-gegen-fakes-gemeinsam-fuer-eine-starke-demokratie-ergebnisse-der-evaluation-des-buergerrats-und-der-online-beteiligung

Bertelsmann Stiftung. (2025b). *Staatliche Maßnahmen gegen Desinformation: Auf dem Weg zu einer resilienten und gesunden Öffentlichkeit*. https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/staatliche-massnahmen-gegen-desinformation-auf-dem-weg-zu-einer-resilienten-und-gesunden-oeffentlichkeit

Boender, T. S., Schneider, P. H., Houareau, C., Wehrli, S., Purnat, T. D., Ishizumi, A., Wilhelm, E., Voegeli, C., Wieler, L. H., & Leuker, C. (2023). Establishing infodemic management in Germany: A framework for social listening and integrated analysis to report infodemic insights at the national public health institute. *JMIR Infodemiology*, *3*, Article e43646. https://doi.org/10.2196/43646

Boschiero, M. N., Palamim, C. V. C., & Marson, F. A. L. (2021). Covid-19 vaccination on Brazil and the crocodile side-effect. *Ethics, Medicine and Public Health*, *17*, Article 100654. https://doi.org/10.1016/j.jemep.2021.100654

Brazilian Ministry of Health. (2025). *Saúde com ciência*. https://www.gov.br/saude/pt-br/assuntos/saude-com-ciencia

Bundesnetzagentur. (2025). *Koordinierungsstelle für digitale Dienste (DSC): Tätigkeitsbericht 2024—Nach § 17 DDG, Art. 55 Verordnung (EU) 2065/2022*. https://www.dsc.bund.de/DSC/DE/Aktuelles/Downloads/DSC_Bericht2024.pdf?__blob=publicationFile&v=6

Burger, J., & Gochfeld, M. (2019). Involving community members in preparedness and resiliency involves bi-directional and iterative communication and actions: A case study of vulnerable populations in New Jersey following Superstorm Sandy. *Journal of Risk Research*, *23*(4), 541–556. https://doi.org/10.1080/13669877.2019.1577590

Cardenuto, J. P., Yang, J., Padilha, R., Wan, R., Moreira, D., Li, H., Wang, S., Andaló, F., Marcel, S., & Rocha, A. (2023). The age of synthetic realities: Challenges and opportunities. *APSIPA Transactions on Signal and Information Processing*, *12*(1), Article e44. https://doi.org/10.1561/116.00000138

Cazzamatta, R., & Sarısakaloğlu, A. (2025). AI-generated misinformation: A case study on emerging trends in fact-checking practices across Brazil, Germany, and the United Kingdom. *Emerging Media*, *3*(2), 214–251. https://doi.org/10.1177/27523543251344971

Chu, H., & Liu, S. (2025). Generating targeted and tailored health communication narratives with AI. *Risk Analysis*, *45*(11), 3505–3518. https://doi.org/10.1111/risa.70076

Chuai, Y., Tian, H., Pröllochs, N., & Lenzini, G. (2024). Did the roll-out of Community Notes reduce engagement with misinformation on X/Twitter? *Proceedings of the ACM on Human-Computer Interaction*, *8*(CSCW2), Article 428. https://doi.org/10.1145/3686967

Conselho Nacional de Ciência e Tecnologia. (2024). *Resolução CCT nº 4, de 08.11.2024*. https://antigo.mctic.gov.br/mctic/opencms/legislacao/outros_atos/resolucoes/Resolucao_CCT_n_4_de_08112024.html

Constituição da República Federativa do Brasil, 1988. https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm

Dalglish, S. L., Khalid, H., & McMahon, S. A. (2020). Document analysis in health policy research: The READ approach. *Health Policy and Planning*, *35*(10), 1424–1431. https://doi.org/10.1093/heapol/czaa064

Dandara, L. (2022, June 8). Cinco dias de fúria: Revolta da Vacina envolveu muito mais do que insatisfação com a vacinação. *Fiocruz*. https://fiocruz.br/noticia/2022/06/cinco-dias-de-furia-revolta-da-vacina-envolveu-muito-mais-do-que-insatisfacao-com

Eisenman, D. P., Cordasco, K. M., Asch, S., Golden, J. F., & Glik, D. (2007). Disaster planning and risk communication with vulnerable communities: Lessons from Hurricane Katrina. *American Journal of Public Health*, *97*(Suppl. 1), S109–S115. https://doi.org/10.2105/AJPH.2005.084335

European Center for Constitutional and Human Rights. (2025). *Hard law/soft law*. https://www.ecchr.eu/en/glossary/hard-law-soft-law

European Commission. (2018). *Tackling online disinformation: A European approach*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52018DC0236

European Commission. (2023, December 18). *Commission opens formal proceedings against X under the Digital Services Act* [Press release]. https://ec.europa.eu/commission/presscorner/detail/en/ip_23_6709

European Commission. (2025a). *Supervision of the designated very large online platforms and search engines under DSA*. https://digital-strategy.ec.europa.eu/en/policies/list-designated-vlops-and-vloses

European Commission. (2025b). *The Code of Conduct on Disinformation*. https://digital-strategy.ec.europa.eu/en/library/code-conduct-disinformation

European Monitoring Centre for Drugs and Drug Addiction. (2023). *Health risk communication strategies for drug checking services*.

Fonseca, B. (2023, May 9). Google pagou mais de meio milhão de reais em anúncios no Facebook contra PL das fake news. *Pública*. https://apublica.org/2023/05/google-pagou-mais-de-meio-milhao-de-reais-em-anuncios-no-facebook-contra-pl-das-fake-news

Frosio, G. (2024). *From the e-Commerce Directive to the Digital Services Act*. *SSRN*. https://doi.org/10.2139/ssrn.4914816

Glik, D. C. (2007). Risk communication for public health emergencies. *Annual Review of Public Health*, *28*, 33–54. https://doi.org/10.1146/annurev.publhealth.28.021406.144123

Graves, L. (2025, January 13). Will the EU fight for the truth on Facebook and Instagram? *The Guardian*. https://www.theguardian.com/technology/2025/jan/13/meta-facebook-factchecking-eu

Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review*, *61*(4), 5–14. https://doi.org/10.1177/0008125619864925

Husovec, M. (2023). Rising above liability: The Digital Service Act as a blueprint for the second generation of global internet rules. *Berkeley Technology Law Journal*, *38*(3), 883–920.

Huxhold, O., Hees, E., & Webster, N. J. (2020). Towards bridging the grey digital divide: Changes in internet access and its predictors from 2002 to 2014 in Germany. *European Journal of Ageing*, *17*(3), 271–280. https://doi.org/10.1007/s10433-020-00552-z

Instituto Democracia em Xeque. (2025). *Democracia em Xeque Institute*. https://en.institutodx.org

International Fact-Checking Network. (2025). *IFCN Code of Principles*. https://ifcncodeofprinciples.poynter.org

Jahangir, R. (2025a, February 25). The EU's Code of Practice on Disinformation is now part of the Digital Services Act. What does it mean? *Tech Policy Press*. https://www.techpolicy.press/the-eus-code-of-practice-on-disinformation-is-now-part-of-the-digital-services-act-what-does-it-mean

Jahangir, R. (2025b, April 5). Understanding the EU's Digital Services Act enforcement against X. *Tech Policy Press*. https://www.techpolicy.press/understanding-the-eus-digital-services-act-enforcement-against-x

Kahn, G. (2025, January 17). Amid war, vicious attacks and political turmoil, global fact-checkers fear the impact of the end of Meta's programme. *Reuters Institute*. https://reutersinstitute.politics.ox.ac.uk/news/amid-war-vicious-attacks-and-political-turmoil-global-fact-checkers-fear-impact-end-metas

Kwek, A. (2017). The indispensability of labelled groups to vulnerability in bioethics. *Bioethics*, *31*(9), 674–682. https://doi.org/10.1111/bioe.12379

Lelo, T. (2022). The rise of the Brazilian fact-checking movement: Between economic sustainability and editorial independence. *Journalism Studies*, *23*(9), 1077–1095. https://doi.org/10.1080/1461670X.2022.2069588

Lupton, D. (1999). *Risk*. Routledge.

McGowan, B. S. (2022). World Health Organization's early AI-supported response with social listening platform. *Journal of the Medical Library Association*, *110*(2), 273–275. https://doi.org/10.5195/jmla.2022.1398

Menezes, L. F., & Mangabeira, M. (2025, August 21). Mensagens reveladas pela PF mostram aliança entre Bolsonaro e EUA contra regulação das 'big techs.' *Aos Fatos*. https://www.aosfatos.org/noticias/mensagens-pf-alianca-bolsonaro-eua-contra-regulacao-big-techs

Meseret, E. (2024, February 8). A global rise in government-led fact-checking initiatives cause concern, worries of misuse. *Poynter*. https://www.poynter.org/ifcn/2024/government-fact-checking-initiatives-worldwide

Miller, M. R., Sehat, C. M., & Jennings, R. (2024). Leveraging AI for public health communication: Opportunities and risks. *Journal of Public Health Management and Practice*, *30*(4), 616–618. https://doi.org/10.1097/PHH.0000000000001986

Monari, A. C. P. (2024). *Quem consome fake news consome fake news? Os usos e sentidos atribuídos à ciência e ao jornalismo em canais sobre Covid-19 no Telegram* [Unpublished doctoral dissertation]. Oswaldo Cruz Foundation.

Moreira, S. V. (2017). Análise documental como método e como técnica. In J. Duarte & A. Barros (Eds.), *Métodos e técnicas de pesquisa em comunicação* (pp. 269–279). Atlas.

Nicholson, P. J. (1999). Communicating health risk. *Occupational Medicine*, *49*(4), 253–256. https://doi.org/10.1093/occmed/49.4.253

Núcleo de Informação e Coordenação do Ponto BR. (2024). *Conectividade significativa: Propostas para medição e o retrato da população no Brasil*.

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

Ogie, R. I., Rho, J. C., & Clarke, R. J. (2018). Artificial intelligence in disaster risk communication: A systematic literature review. In *5th International Conference on Information and Communication Technologies for Disaster Management (ICT-DM)* (pp. 1–8). IEEE. https://doi.org/10.1109/ICT-DM.2018.8636380

Ott, T. (2025). Invisibility and exposure: Challenges of using AI in healthcare. *Zeitschrift für medizinische Ethik*, *71*(4), 488–499. https://doi.org/10.30965/29498570-20250143

Ozawa, J. V. S., Woolley, S. C., Straubhaar, J., Riedl, M. J., Joseff, K., & Gursky, J. (2023). How disinformation on WhatsApp went from campaign weapon to governmental propaganda in Brazil. *Social Media + Society*, *9*(1). https://doi.org/10.1177/20563051231160632

Pacheco, D. (2023, July 28). Navegar é preciso! Regular (as redes) também. *Jornal da USP*. https://jornal.usp.br/atualidades/especial-desconstruindo-a-desinformacao-navegar-e-preciso-regular-as-redes-tambem

Peters, E., Lipkus, I., & Diefenbach, M. A. (2006). The functions of affect in health communications and in the construction of health preferences. *Journal of Communication*, *56*(Suppl. 1), S140–S162. https://doi.org/10.1111/j.1460-2466.2006.00287.x

Presidência da República do Brasil. (1990). *Lei nº 8.080, de 19 de Setembro de 1990*. https://www.planalto.gov.br/ccivil_03/leis/l8080.htm

Presidência da República do Brasil. (2014). *Lei nº 12.965, de 23 de Abril de 2014*. https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm

Purnat, T. D., Ishizumi, A., Yau, B., White, B., Bertrand-Ferrandis, C., Briand, S., & Nguyen, T. (2022). Delivering actionable infodemic insights and recommendations for the Covid-19 pandemic response. *European Journal of Public Health*, *32*(Suppl. 3), Article ckac129.645. https://doi.org/10.1093/eurpub/ckac129.645

Rede Nacional de Combate à Desinformação. (2025). *Quem somos*. https://rncd.org/sobre-nos

Reynolds, B., & Seeger, M. W. (2005). Crisis and emergency risk communication as an integrative model. *Journal of Health Communication*, *10*(1), 43–55. https://doi.org/10.1080/10810730590904571

Ricard, J., & Medeiros, J. (2020, April 17). Using misinformation as a political weapon: Covid-19 and Bolsonaro in Brazil. *Harvard Kennedy School Misinformation Review*. https://doi.org/10.37016/mr-2020-013

Ruhrmann, G., & Guenther, L. (2025). Risikokommunikation. In Bundeszentrale für gesundheitliche Aufklärung (Ed.), *Leitbegriffe der Gesundheitsförderung und Prävention. Glossar zu Konzepten, Strategien und Methoden*. https://doi.org/10.17623/BIOEG:Q4-i103-2.0

Sacramento, I., de Albuquerque, A., Araújo, K., Paes, A., Seixas, F., Gomes, L. G., Cunha, S. E., Cardoso, N. d. O., Pinto, P., & Moreira, T. (2024). *Guia para profissionais de saúde. Desinformação sobre saúde: Vamos enfrentar esse problema?* Ministério da Saúde.

Saurwein, F., & Spencer-Smith, C. (2020). Combating disinformation on social media: Multilevel governance and distributed accountability in Europe. *Digital Journalism*, *8*(6), 820–841. https://doi.org/10.1080/21670811.2020.1765401

Senado Federal do Brasil. (2023). *Projeto de Lei nº 2338, de 2023*. https://www25.senado.leg.br/web/atividade/materias/-/materia/157233

STF define parâmetros para responsabilização de plataformas por conteúdos de terceiros. (2025, June 26). *Supremo Tribunal Federal*. https://noticias.stf.jus.br/postsnoticias/stf-define-parametros-para-responsabilizacao-de-plataformas-por-conteudos-de-terceiros

Tabak, B. M., Froner, M. B., Cardoso, D. H. R., da Conceição, L. A., & de Almeida, R. D. C. (2025). Understanding digital health literacy in Brazil: Findings from the eHEALS survey. *Public Health*, *246*, Article 105828. https://doi.org/10.1016/j.puhe.2025.105828

Teo, M., Goonetilleke, A., Ahankoob, A., Deilami, K., & Lawie, M. (2018). Disaster awareness and information seeking behaviour among residents from low socio-economic backgrounds. *International Journal of Disaster Risk Reduction*, *31*, 1121–1131. https://doi.org/10.1016/j.ijdrr.2018.09.008

UNICEF. (2020). *Covid-19: How to include marginalized and vulnerable people in risk communication and community engagement*. http://data.unhcr.org/en/documents/details/75291

Waisbord, S. (2018). Truth is what happens to news: On journalism, fake news, and post-truth. *Journalism Studies*, *19*(13), 1866–1878. https://doi.org/10.1080/1461670X.2018.1492881

Waisbord, S. (2020). Fake health news in the new regime of truth and (mis)information. *RECIIS*, *14*(1). https://doi.org/10.29397/reciis.v14i1.1953

Wardle, C., & Derakhshan, H. (2017). *Information disorder*. Council of Europe.

Winarnita, M., Yang, F., Saldin, M., Weng, E., Haw, A., Nilan, P., & Rastiya, A. (2025). Indonesian–Australian media during an infodemic: Fostering trust and social resilience through translation as care. *Media International Australia*. Advance online publication. https://doi.org/10.1177/1329878X251342620

World Bank. (2024). *Individuals using the internet (% of population)—Brazil*. https://data.worldbank.org/indicator/IT.NET.USER.ZS?locations=BR

World Health Organization. (2018). *Communicating risk in public health emergencies: A WHO guideline for emergency risk communication (ERC) policy and practice*. https://www.who.int/publications/i/item/9789241550208

World Health Organization. (2025). *Infodemic*. https://www.who.int/health-topics/infodemic#tab=tab_1

Zingaretti, N., Grapini, M., Moretti, A., Noichl, M., Picierno, P., Benifei, B., Incir, E., Saliba, A. A., Gualmini, E., Schieder, A., Glucksmann, R., Laureti, C., Nemec, M., Cereza, L. B., Grossmann, E., Ceulemans, E., Jouvet, P., Laurent, M., Ridel, C., . . . Ruotolo, S. (2025). *Parliamentary question: Meta's recent decision to end fact-checking on Facebook, Instagram and WhatsApp in the United States*. https://www.europarl.europa.eu/doceo/document/E-10-2025-000423_EN.html#def1

About the Authors

Ana Carolina Monari (PhD) is a postdoctoral researcher at the Artificial Intelligence Laboratory (Recod.ai), University of Campinas, Brazil.

Tabea Ott (Ass.-Prof. Dr.) is an assistant professor of medical ethics and diaconal sciences at the University of Vienna, Austria. Previously, she was a postdoctoral researcher at Friedrich-Alexander-Universität Erlangen-Nürnberg.