



CADERNOS DE DEREITO ACTUAL

[www.cadernosdedereitoactual.es](http://www.cadernosdedereitoactual.es)

© **Cadernos de Direito Actual** Nº 25. Núm. Extraordinario (2024), pp. 75-94  
·ISSN 2340-860X - ·ISSNe 2386-5229

## **Disinformation, digital platforms and digital human rights** *Desinformação, plataformas digitais e direitos humanos digitais*

**Raphael Rodrigues Valença de Oliveira**<sup>1</sup>

*Federal University of Rio Grande do Norte*

**Marco Bruno Miranda Clementino**<sup>2</sup>

*Federal University of Rio Grande do Norte*

**Summary:** 1. Introduction. 2. The Informational Order and digital human rights. 2.1. The Informational Order in time and space. 2.2. Disinformation as an unwanted product of the Informational Order and its legal repercussions. 3. Disinformation as a transnational risk to the IAHRs. 4. Conclusions. References.

**Abstract:** The article analyzes the legal repercussions of the spread of disinformation practiced through digital platforms in light of the rights and obligations imposed by the Inter-American System for the Protection of Human Rights (IACHR). It also seeks to contextualize the current moment of the Informational Order and its relationship with digital human rights, in order to understand the impacts of disinformation perpetrated through digital platforms as a transnational risk to the IACHR. Finally, solutions are presented to mitigate the negative impact of disinformation on the promotion and guarantee of the human rights set out in the American Convention on Human Rights (ACHR). A bibliographical and documentary survey is carried out, with emphasis on national and foreign doctrinal, legislative and jurisprudential sources.

**Keywords:** Disinformation. Digital Platforms. Digital Human Rights. Inter-American System for the Protection of Human Rights.

**Resumo:** O presente trabalho analisa as repercussões jurídicas da disseminação da desinformação praticada por meio das plataformas digitais à luz dos direitos e obrigações impostos pelo Sistema Interamericano de Proteção aos Direitos Humanos (SIDH). Outrossim, busca-se contextualizar o atual momento da Ordem Informacional e a sua relação com os direitos humanos digitais, visando compreender os impactos da desinformação perpetrada por meio das plataformas digitais como risco transnacional ao SIDH. Ao final, apresentam-se soluções objetivando mitigar o impacto negativo da desinformação na promoção e garantia dos direitos humanos previstos na Convenção Americana sobre Direitos Humanos (CADH). Realiza-se levantamento bibliográfico e

<sup>1</sup> Master's Degree in Law from the PPGD at the Federal University of Rio Grande do Norte. Attorney for the Union – Attorney General's Office. E-mail: [raphael.oliveira@agu.gov.br](mailto:raphael.oliveira@agu.gov.br).

<sup>2</sup> Professor at the Federal University of Rio Grande do Norte. Master and Doctor in Law. Federal Judge. Email: [marcobrunomiranda@gmail.com](mailto:marcobrunomiranda@gmail.com).

documental, com ênfase em fontes doutrinárias, legislativas e jurisprudenciais, nacionais e estrangeiras.

**Palavras-chave:** Desinformação. Plataformas Digitais. Direitos Humanos Digitais. Sistema Interamericano de Proteção aos Direitos Humanos.

## 1 INTRODUCTION

The Informational Order brought with it a revolution in the way information is produced, disseminated and consumed by digital subjects. Digital platforms, such as social media/networks and search engines, with the consolidation of the digitalization of private life and the virtualization of the intimacy of the digital citizen, have become unprecedented central spaces in the formation of public opinion and the construction of democratic debate. However, umbilically associated with the undeniable benefits of these technologies, proportional challenges have emerged, among which the proliferation of disinformation stands out. This phenomenon, characterized by the deliberate dissemination of false or misleading information, presents substantial risks to society, especially with regard to the protection of human rights enshrined in the Inter-American System for the Protection of Human Rights (IAHRS), the normative focus of this research.

In this step, it is verified that the practice of disinformation, especially when consummated through digital networks, can vilify a series of rights that have been positivized in the inter-American sphere, such as the rights to information, personal integrity, protection of honor and dignity, freedom of conscience and religion, freedom of thought and expression, among others, recorded in the American Convention on Human Rights (ACHR or American Convention), which will be subject a specific examination.

In addition, disinformation can exacerbate the violation of civil and political rights in the strict sense, including the manipulation of electoral processes and the undermining of democracy. In crisis contexts, such as pandemics or armed conflicts, disinformation can also threaten the right to health and life by spreading dangerous rumours or discouraging preventive behaviour.

In this context, this article seeks to explore the risks that disinformation perpetrated through digital platforms poses to human rights. To this end, the main forms of disinformation, the mechanisms used for its propagation, and the direct and indirect impacts on various digital human rights will be analyzed, that is, traditional human rights that now represent autonomous legal assets in the online environment.

By addressing these issues, the work intends to contribute to the construction of a legal framework that, with due scientific rigor, allows the confrontation of the implications of digital disinformation and, simultaneously, formulate effective strategies to protect human rights in an increasingly complex and interconnected informational environment.

## 2 THE INFORMATIONAL ORDER AND DIGITAL HUMAN RIGHTS

### 2.1. The Informational Order in time and space

From the second half of the twentieth century, with the advent of the so-called post-industrial society, precursor of another stage of the technological revolution that was approaching, the human being began a process of expanding the *territorial locus* (historically attached to his village, tribe, fiefdom, town and city), detaching himself from the land as a static and limiting element of human experience, starting to interact in an innovative, dynamic way (social relationships become more intense and closer, as a result of more direct contact) and now with different actors of the collective (such as people other than those in their close life, companies from other cities and countries, states, international organizations, among others).

This rupture in the way people related, lived and worked, announced a new Order, which would not only impact the most different fields of science (sociology, economics, law, among others), but would also profoundly re-signify the conceptions of the world, order, society, capitalism, State, constitutionalism and the very meaning of being human.

This is the so-called Informational Order<sup>3</sup>, built from the fusion of some distinctive landmarks.<sup>4</sup>

It should be noted, from the outset, that the phenomena discussed below did not occur in isolation, apart from the others, much less did they follow a chronological order of sequenced factual concretion, revealing, therefore, with greater rigor, relations of complementarity and coordination than cause and effect itself, although, with the exception, when so inferred, the due reservations will be made in the text.

The first symbol that demarcated the change in the logic thus projected is the advent or consolidation of globalization, interpreted by the doctrine<sup>5</sup> as a movement of reduction of spatial distances in the world, guided by the perspective of convergence of territorial nuclei, people, goods and services, which, in addition to not being circumscribed to the traditional borders demarcated by geopolitics, nor does it submit to the national wills of sovereign states. It is understood as a manifestation that reshapes the world order, imposing new roles, duties and challenges on the traditional subjects of international law. As will be demonstrated, globalization is divided into cycles or phases, enjoying, at the moment, effects that have never been observed before, a magnitude justified, according to the hypotheses of this research, due to the unprecedented coexistence with the next milestones detected.

This new cycle or current moment of globalization was unimaginable by the common citizen of the 1960s and 1970s, since, regardless of their place of residence, they had little interaction with friends or family members who lived in a neighboring state, regardless of the degree of intimacy that may have existed between them. Depending on the country, it would be difficult to build bonds of friendship or maintain close blood ties with those who lived beyond their city, precisely because of the difficulties inherent in distance communication, as well as face-to-face contact. This social model of interaction between individuals has been exponentially remodeled over the last 60 (sixty years), a chronological interval marked by the doctrine as the emergence of a new classification of society.

The doctrine is not peaceful regarding the concept of this new model of social organization, concentrating several terminologies<sup>6</sup> to illustrate this new arrangement resulting from the social contract signed by individuals, which seem to translate different perspectives on the same phenomenon investigated. In this sense, for the purposes of this research, the concept of information society will be adopted to identify this new way of relating socially, responsible for elevating information, knowledge and data to protagonists of the connections between individuals, redesigning the priorities as well as the focuses of interest of the subjects.

The emergence of this social model is a product of the gradual and historical displacement of the dominant economic structure, once based on the land (agrarian society) and the factory (industrial society), to today take root in information. The doctrine points out that it was not the information itself that evoked this protagonism, but the evolution of information and communication technologies (ICTs) that were responsible for

---

<sup>3</sup> For those wishing to study the concept of Informational Order in greater depth, see: OLIVEIRA, R.R.V. *Regime brasileiro de transferência internacional de dados pessoais*, Lumen Juris, Rio de Janeiro, 2024.

<sup>4</sup> OLIVEIRA, R.R.V; CLEMENTINO, M.B.M. "Os MLATS vigentes no Brasil e as implicações na soberania contemporânea: internacionalização, territorialidade e jurisdição", *Revista FIDES*, v. 12, n. 1, 2021, p. 830-849.

<sup>5</sup> VON BOGDANDY, A. "Globalization and Europe: How to Square Democracy, Globalization, and International Law". *European Journal of International Law*, v. 15, n. 5, 2004, p. 885-906, <https://doi.org/10.1093/ejil/15.5.885>. Available in <https://academic.oup.com/ejil/article/15/5/885/533488>.

<sup>6</sup> The expressions most used by literature, whether national or foreign, are: information society, knowledge society, post-industrial society and digital society. Notwithstanding the relevance of the classifications adopted, considering that for the purposes of the present research, they will not bring effects different from those provided by the terminology "information society", this was adopted as the chosen one.

allowing and accelerating this process, still in progress of construction of the current social structure<sup>7</sup>.

The nomenclature "information society" is not exactly new, nor is it immune to criticism, as pointed out, especially when opposed by authors who prefer to label the current moment as a "knowledge society", "post-industrial society" and "digital society". In essence, despite the preference for one terminology or another, it is found that the information society can be conceived as a social nucleus that incorporates information and communication technology (ICT) as a fundamental mechanism for the construction of the production chain, centralizing information, knowledge and data as core resources of human relationships. It is, however, an organizational context that would be re-signified again after the popularization of the internet.

With the spread of the World Wide Web and the expansion of commercial services, information communication networks have been improving and gaining even greater popularity. At this moment, which dates back to the final decade of the twentieth century and the beginning of the twenty-first century, with greater or lesser intensity depending on the degree of technological development of the country under examination, the network infrastructure and the available applications began to harmonize, providing another degree of experience to users, a scenario that would be inviting not only for natural people, but also for private companies, state entities and international organizations to launch themselves into the digital society<sup>8</sup>.

This new element boosted the instantaneous and daily contact between individuals, who began to carry out various operations with data on a daily basis, especially universalizing the access and sharing of (massified) information, transforming the large network into a new space for the promotion of rights, freedoms, knowledge, social interaction, as well as the production of wealth. However, not only values dear and desired by the social, legal and economic orders were fostered. The internet has transformed production processes, attracting with it to the digital *locus* several of the problems associated with and resulting from physical interactions, such as market disputes, conflicts of interest, in addition to other challenges that are inherent to the human condition in a collective context, such as crimes, misdemeanors, among other unwanted conducts that are the object of disapproval by the law.

The combination of these two elements, information society and popularization of the internet, further accentuated this profile of the new collective organization, highlighting the role of data as the central axis of the most diverse orders, a framework that inspired the development of several concepts intended to differentiate these social (data-driven society), economic (data-driven economy) and legal (legal informational order).

In addition to the causes already mentioned, another impactful movement claims prominence in this informational revolution: the advent of the Fourth Industrial Revolution<sup>9</sup>, conceived as a new stage of industrial transformation, characterized by the development of technologies in areas such as genetics, physics, robotics, microelectronics and other niches related to digital innovations<sup>10</sup>. Regardless of the adherence to certain terminologies, it is easy to see that signs such as artificial intelligence, machine learning, the internet of things, blockchain, and virtual reality, in addition to their incorporation into everyday life, have reshaped the way individuals live socially, their behavior patterns, as well as the very meaning of being human.

---

<sup>7</sup> BELL, D. *The Coming of Post-Industrial Society: A Venture in Social Forecasting*, Basic Books, New York, 1973.

<sup>8</sup> About the history and development of the internet in the world and in Brazil, see: CARVALHO, M. S. R. M. *A trajetória da Internet no Brasil: do surgimento das redes de computadores à instituição dos mecanismos de governança*. Dissertação de mestrado, Programa de Pós-Graduação em Engenharia de Sistemas de Computação, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2006.

<sup>9</sup> SCHWAB, K. *The Fourth Industrial Revolution*, Penguin Books, London, 2017; SCHWAB, K.; DAVIS, N. *Shaping the Fourth Industrial Revolution*, World Economic Forum, Geneva, 2018.

<sup>10</sup> VERONESE, A. "The fourth industrial revolution and blockchain: social values and trust". *ADENAUER NOTEBOOKS XXI*, Konrad Adenauer Foundation, Rio de Janeiro, 2020.

On the economic level, the repercussions were so substantial that they even impacted the very conception of the capitalist economic system, prevalent since the collapse of the former Soviet Union, allowing space for the development of the so-called Surveillance Capitalism<sup>11</sup>, identified as a particular species of the capitalism genus, based on the monetization of data acquired by surveillance, that is, An economic platform that considers human experience as pure material (raw data) and free for hidden business practices of extracting, sharing, predicting and selling data<sup>12</sup>. In this sense, if even raw data, which are those that have not yet been processed, already have high monetary repercussions, since they infer necessary, useful or even vital information for the performance of economic activity, with due economic treatment, the capacity to produce wealth then reaches levels that are difficult to measure, logic that, by itself, would be enough to demonstrate the risks to which the holders of these infinite inputs are subjected.

In this step, given the impressive technological revolution observed in fields such as machine learning, artificial intelligence, the internet of things, and virtual reality, combined with a natural rooting of these mechanisms in the most intimate routines of collective and individual action, a scenario of inevitable indexation<sup>13</sup> has been structured and digital tracking (online tracking), in which monitoring and hyper-surveillance have been revealed as common practices and even in a way fostered by companies and States, under the most diverse reasons<sup>14</sup>.

The first reason is based on the economic reflex, because if personal data, in this data-driven economy, has become the inputs that generate value, it is evident that the production system will adapt to extract the highest yield (benefit) from them, at the lowest possible cost, which is why thousands of examples of companies offering various benefits to users are seen daily. Consumers and even third parties apparently uninterested in evaluating a service, answering a survey or commenting on a certain location, regardless of whether they have actually contracted the service, know about the product to be asked or know the place commented, since the data processing method in surveillance capitalism is supported by mass collection, wholesale (big data). Thus, due to the high volume of data collected (a kind of treatment) in this first level of extraction, the concern is only the generation of data for data's sake, feeding back into an economic system that demands data in order to reach its full (and immeasurable) potential. Therefore, any hypothetical judgment of veracity and falsity is not important at this stage for the algorithm.

The second aspect lies in a consequentialist analysis, based on the cost-benefit perspective. The benefit of the unrestricted collection of data from the holders, for the most part, without any type of prior consent, counterpart or even knowledge<sup>15</sup> of it, means free raw material available for appropriation without any kind of difficulty. The cost, in turn, would be contemplated if the protective legislation were effective, both so that those responsible for data processing (controller and operator) ensure compliance with their duties and obligations, as well as so that, in cases of non-compliance with these commandments, they would be held responsible in a manner proportional to the seriousness of the infractions and irregularities, a reality that is not yet applicable in Brazil, for example, as will be demonstrated.

The third point concerns the fact that these practices have been fostered by States, transnational corporations and other major players in the information market to create a large set of data that only mirrors one side of the faces (one-way mirror), since only citizens, users and consumers are the ones who supply this big data, providing the

<sup>11</sup> ZUBOFF, S. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, Profile Books, London, 2019.

<sup>12</sup> This social dynamic was also masterfully examined by David Lyon, in his work *Surveillance society: monitoring everyday life*, Open University Press, Buckingham, 2001.

<sup>13</sup> INTRONA, L.; NISSEBAM, H. "Shaping the web: why the politics of search engines matters", *The Information Society*, v. 16, 2000, p. 169-185.

<sup>14</sup> HOOFNAGLE, C. J.; SOLTANI, A.; GOOD, N.; WANBACH, D. J. "Behavioral advertising: the offer you can't refuse", *Harvard Law & Policy Review*, v.6, 2012, p. 273-296. Available in: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2137601](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2137601).

<sup>15</sup> ZUBOFF, S. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, Profile Books, London, 2019, p.17-25.

respective controllers (States and large companies, mainly) with unimaginable knowledge, control and power over the surveilled<sup>16</sup>, who in addition to not knowing that they are being monitored, are unaware of the data that is being collected and for what purposes they are being processed, failing to comply with several precepts of specific protective laws, as well as fundamental rights and guarantees that are basic in most of the legal systems minimally structured in the world.

Such abuses are sources of several obscure practices, such as profiling, through which States, companies and subjects form identification profiles of citizens, users and consumers (data subjects), associating preferences, financial standard, adopted routes, among other particular characteristics, aiming to predict desired behaviors or avoid unwanted ones, in order to obtain advantages. A famous precedent for this hidden practice occurred in the Cambridge Analytica case, in which the political consulting firm responsible for advising then-candidate Donald Trump improperly used data from more than 87 (eighty-seven) million users of the social network Facebook, collecting them through an apparently harmless application accessed within the platform of the American giant, with the intention not only of building user profiles, but also of manipulating them to vote for Trump.

Finally, a practice that also represents an important behavioral change that did not exist before the advent of social networks deserves special mention, a concrete product of a process identified in this research as virtualization of intimacy and digitalization of private life<sup>17</sup>, in which, deliberately, people choose, through automatic volitional manifestations, without critical judgment on the resulting repercussions, make various sensitive personal data available on global platforms that operate millions of data in seconds<sup>18</sup>, offering access, use, reproduction and storage, for example, of this data to people physically located in the most diverse countries, subject, therefore, to the effects of multiple legal orders.

As a result, globalization, the information society, the popularization of the internet, the data-driven economy, surveillance capitalism, the industrial revolution, the virtualization of intimacy and the digitalization of private life call for the role of protagonists of this new Informational Order that is being built by large and fierce steps, propelling information and human rights to the epicenter of this digital transformation.

## **2.2. Disinformation as an unwanted product of the Informational Order and its legal repercussions**

In a society in which information has a prominent seat, it is necessary to understand the reasons that justify the constant disputes to manipulate it and, above all, control it, because, as previously demonstrated, mere access to information represents today a valuable input in the Informational Order. At this stage, although it is undeniable that the digital age has provided significant advances in the expansion of democratic quality, solidified above all in the plurality of sources of information available, as well as in the volume, speed and viral capacity that this information can reach<sup>19</sup> through the informational circulation sponsored via digital platforms, on the other hand, there is also an expansion of the information that can be achieved through the dissemination of information and the spread of information through digital platforms. concerns about the

<sup>16</sup> PASQUALE, F. *The black box society. The secret algorithms that control money and information*, Harvard University Press, Cambridge, 2015, p. 9.

<sup>17</sup> In this sense, some authors claim that to exist is to be indexed in a search engine. INTRONA, L; NISSEBAM, H. "Shaping the web: why the politics of search engines matters", *The Information Society*, v. 16, 2000, p. 169-185.

<sup>18</sup> GANDOMI, A; MURTAZA, H. "Beyond the hype: Big data concepts, methods, and analytics", *International Journal of Information Management*, v. 35, n. 2, 2015, p.137-144, ISSN 0268-4012, <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>. Available in:

<https://www.sciencedirect.com/science/article/pii/S0268401214001066>. "Beaver, Kumar, Li, Sobel, and Vajgel (2010) report that Facebook processes up to one million photographs per second".

<sup>19</sup> VOSOUGHI, S; ROY, D; ARAL, S. "The Spread of True and False News Online", 2018. Available in: <https://science.sciencemag.org/content/359/6380/1146>.

dysfunctions that can affect this informational ecosystem, concerns that inspire and demand special care on the part of States and the international community.

In this wake, some threats to the basic properties of information security<sup>20</sup> have received special prominence, whether those that threaten the availability of information<sup>21</sup>, such as censorship and authoritarianism, or those that are even more controversial today, such as those that threaten the integrity of information, such as disinformation, hate speech and false information. Considering the impact and harmful reach, it is noteworthy that the present research will be dedicated to the first two forms: disinformation and hate speech.

Using, initially, the contributions made by the ISHR itself, we extract the definition presented by the Guide to guarantee freedom of expression in the face of deliberate disinformation in electoral contexts<sup>22</sup>, from the Office of the Special Rapporteur for Freedom of Expression of the Inter-American Commission on Human Rights (IACHR), which conceives disinformation as the massive dissemination of false information (a) with the intention of misleading the public and (b) with awareness of its falsity.

In 2018, the European Union, in turn, through a Report promoted by the High Level Group on Fake News and Online Disinformation, also tried to conceptualize disinformation, understanding that disinformation as defined in this Report includes all forms of false, inaccurate, or misleading information designed, presented and promoted to intentionally cause public harm or for profit.<sup>23</sup>

In a complementary sense, the United Nations Report on Information Integrity (2023), referring to a concept presented by UNESCO, presents disinformation as false or misleading content that can cause specific harm, regardless of motivations, awareness, or behaviors.<sup>24</sup>

From these definitions, common elements capable of identifying the concept of disinformation stand out: (i) content known to be false and (ii) created/disseminated with the intention of deceiving the public/causing harm.

Hate speech, in turn, another serious threat to the integrity of information, according to the working definition of the United Nations Strategy and Plan of Action on Hate Speech<sup>25</sup>, is "any type of oral, written or behavioral communication, which attacks or uses pejorative or discriminatory language with reference to a person or group based on who they are, that is, based on their religion, ethnicity, nationality, race, color, descent, gender or other identity factor".

In the context of the IAHRs, the apology for hatred that constitutes incitement to violence is one of the speeches not covered by freedom of expression, under the terms of Article 13.5 of the ACHR, which expressly provides that "the law must prohibit all propaganda in favor of war, as well as all apology for national, racial or religious hatred that constitutes incitement to discrimination, to hostility, crime or violence".

Having elaborated the necessary premises to conceptualize disinformation and hate speech, we now enter the legal repercussions of these two products, which have

<sup>20</sup> By way of illustration, the four main basic properties of information security that are internationally recognized are confidentiality, integrity, availability, and authenticity.

<sup>21</sup> Property responsible for ensuring that information is accessible or usable by a duly authorized individual or system or entity.

<sup>22</sup> ORGANIZATION OF AMERICAN STATES (OAS). Available in: [https://www.oas.org/es/cidh/expresion/publicaciones/Guia\\_Desinformacion\\_VF.pdf](https://www.oas.org/es/cidh/expresion/publicaciones/Guia_Desinformacion_VF.pdf).

<sup>23</sup> EUROPEAN COMMISSION, Directorate-General for Communications Networks, Content and Technology, "A multi-dimensional approach to disinformation – Report of the independent High level Group on fake news and online disinformation, Publications Office", 2018, Available in: <https://data.europa.eu/doi/10.2759/739290>.

<sup>24</sup> UNITED NATIONS. "Policy Brief for Our Common Agenda: Information Integrity of Digital Platforms", 2023. Available in: [https://brasil.un.org/sites/default/files/2023-10/ONU\\_Integridade\\_Informacao\\_Plataformas\\_Digitais\\_Informe-Secretario-Geral\\_2023.pdf](https://brasil.un.org/sites/default/files/2023-10/ONU_Integridade_Informacao_Plataformas_Digitais_Informe-Secretario-Geral_2023.pdf).

<sup>25</sup> UNITED NATIONS. "Strategy and plan of action on hate speech", 2019. Available in: [https://www.un.org/en/genocideprevention/documents/advising-and-mobilizing/Action\\_plan\\_on\\_hate\\_speech\\_EN.pdf](https://www.un.org/en/genocideprevention/documents/advising-and-mobilizing/Action_plan_on_hate_speech_EN.pdf).

affronted the regular exercise not only of freedom of expression and the right to information, but also of other human rights and even democratic institutions.

In this step, despite the unequivocal advantages promoted with the possibility that any citizen can produce information/content, without any editorial control (control that only occurs in traditional mass communication vehicles), this dissemination, practically without limits of information, has also brought with it some harmful effects, given the commission of abuses and illicit speeches. A typical example was experienced especially during the COVID-19 pandemic, which led the World Health Organization to adopt the term "infodemic", used to describe "an excess of information, some accurate and some not, that makes it difficult to find reliable sources and reliable guidance when needed".<sup>26</sup>

Illustrating the harmful impacts that disinformation and hate speech can have on the most diverse fields, such as health, the environment, gender equality, vulnerable groups, marginalized minorities, the electoral process and the democratic regime.

As for health, the spread of misinformation represents a serious threat to the public system in general, especially to vaccination campaigns. When inaccurate or false information circulates, barriers to access to health care and adherence to vaccines are created. This can result in serious consequences, such as outbreaks of vaccine-preventable diseases and increased morbidity and mortality. In addition, misinformation can undermine trust in health authorities and science, leading people to make decisions based on misinformation, putting not only their own health but also the health of their communities at risk. This situation is even more serious when it occurs in the face of a crisis, emergency, or conflict.

About the environment and climatic conditions, the misinformation conveyed about states of climate emergency has undermined urgent actions necessary to ensure a habitable present and future on planet Earth, as evidenced by the experience of coping with recent flood events in the South and prolonged droughts in certain regions of the North and Northeast of Brazil.

In relation to gender equality, vulnerable groups and marginalized minorities, these themes and subjects continue to be targets of hate speech, which has resulted in scenarios of social, economic and political exclusion. A common example can be observed when women, blacks, members of the LGBTQIAPN+ community, journalists and representatives of civil society run for office. A group that deserves superior care is children and adolescents, whether from the aspect of their mental and emotional health (with multiplied scenarios of how much misinformation and hate speech can cause or aggravate anxiety or depression), in favor of protection against exploitation and abuse.

With regard to the electoral process and the democratic regime, disinformation has an impact on the formulation of public opinion, from the moment it (a) can mistakenly influence the voter's perception of candidates, political issues and parties; (b) it can create conspiracy theories that divert attention from issues relevant and important to voters; (c) it can erode the confidence of voters in the electoral process, inferring, therefore, that its massive and unlimited dissemination greatly affronts the free and informed vote and, consequently, the electoral process itself in the States.

Considering this context, it is essential, therefore, that measures must be adopted to combat the aforementioned threats to the integrity of information - more precisely, disinformation and hate speech - in order to safeguard human rights, as well as the rule of law itself. With the dysfunctions pointed out, it is to be deduced that the tools currently available have not been able to contain its dissemination - dysfunctions that, it should be repeated, are not covered by freedom of expression. In this wake, and considering international parameters, it is necessary to develop new paths, especially normative, to safeguard the integrity of the information.

Furthermore, progress in the adoption of new medicines seems to become more effective as they are incorporated as a duty not only of States and the international

---

<sup>26</sup> In this regard, see: PAN AMERICAN HEALTH ORGANIZATION. "Understanding the Infodemic and Misinformation in the fight against COVID-19". (<https://iris.paho.org/handle/10665.2/52052>). Available in [https://iris.paho.org/bitstream/handle/10665.2/52054/Factsheet-Infodemic\\_por.pdf?sequence=16](https://iris.paho.org/bitstream/handle/10665.2/52054/Factsheet-Infodemic_por.pdf?sequence=16).



community, but of society, digital platforms, advertisers and all those actors in the communication process that can, in some way, influence the guarantee of a healthy information ecosystem.

### **3 DISINFORMATION AS A TRANSNATIONAL RISK TO THE IAHRs**

As seen, disinformation has proved to be an unwanted product of the Informational Order itself, challenging not only jurists, but all those who live and relate to the information society and who aspire to coexist in a democratic regime, embodied in the promotion of personal freedom and social justice. In addition, specifically, it can be stated that disinformation has had several negative impacts on the promotion and guarantee of human rights in digital environments, notably those provided for in the ACHR.

In this section, we present the comparison of the conventional provisions that attract greater judgment of violation due to the phenomenon of disinformation and hate speech disseminated through large digital platforms. The provisions will be developed in obedience to the order contained in the conventional text.

At the outset, it should be noted that disinformation and hate speech have demonstrated substantial power to attack the right to personal integrity, provided for in article 5 of the ACHR, while it may compromise the physical, psychic and moral integrity of the individual.

Physical integrity may be vilified when misinformation instigates the holder to adopt conducts that violate his bodily fullness and his right to health, such as adherence to medical treatments that are known to be false or devoid of any scientific basis. In addition, even the circulation of information that aims to discourage the adoption of treatments that have been scientifically proven to be effective, discrediting them, has resulted in collective scenarios that have generated and may continue to generate avoidable damage to the health of the population<sup>27</sup>.

In a complementary sense, it should be remembered that, in emergency situations or natural disasters, such as those experienced during the COVID-19 pandemic and the floods that affected the southern region of Brazil<sup>28</sup>, disinformation has demonstrated – and continues to demonstrate – that it enjoys the power to induce people to make decisions that put their own lives in danger, such as the disinformation campaign that associated the application of the COVID-19 virus immunization vaccine with the emergence of diseases collateral, correlation without any scientific support<sup>29</sup>. On this subject, it is important to highlight the position enshrined in the document "Disinformation, pandemic and human rights"<sup>30</sup>, approved by the IACHR on June 15, 2022:

During the COVID-19 pandemic, that began toward the end of 2019 and that took over the world in the years that followed, disinformation once again emerged as a serious problem. If during some of the recent electoral processes disinformation had posed a threat to democratic systems, during the pandemic this risk was associated with State's capacity to design and implement efficient public health responses and the consequences for public health of the dissemination of false guidance and ineffective or even harmful remedies. First of all, protecting public health from a threat like COVID-19

<sup>27</sup> WORLD HEALTH ORGANIZATION. "Disinformation and public health". Available in: <https://www.who.int/news-room/questions-and-answers/item/disinformation-and-public-health>.

<sup>28</sup> BRASIL. RIO GRANDE DO SUL. DECRETO nº. 57,603, de 5 de maio de 2024. Amends Decree No. 57,600, of May 4, 2024, which reiterates the state of public calamity in the territory of the State of Rio Grande do Sul affected by the climatic events of intense rainfall, COBRADE 1.3.2.1.4, which occur in the period from April 24 to May 2024, and specifies the affected Municipalities.

<sup>29</sup> In this sense, see: SKAFLE, I., NORDAHL-HANSEN, A., QUINTANA, D. S., WYNN, R., & GABARRON, E. "Misinformation About COVID-19 Vaccines on Social Media: Rapid Review", *Journal of Medical Internet Research*, 2022. Available in: <https://doi.org/10.2196/37367>; Within the scope of OAS, see: <https://www.oas.org/ext/en/main/covid-19/Repository/PgrID/3434/PageID/28/pid/3434/mcat/3436/acat/1/tagid/1440/tagname/covid-19>

<sup>30</sup> INTER-AMERICAN COMMISSION ON HUMAN RIGHTS. "Disinformation, pandemic and human rights", 2023.

is a legitimate objective from the perspective of inter-American human rights standards. The protection of "public health or morals" is a phrase that numerous of the rights clauses in the American Convention use explicitly and consistently, including Article 13. This legitimate objective does not, however, grant authorities *carte blanche* to do what they want, and no legitimate objective does. Rather, it is a step in the required "tripartite test" that the inter-American system and its organs use to weigh restrictions on the human rights they are bound to protect. Also, any restriction must be established formally and materially by law,<sup>19</sup> and the restriction must be necessary in a democratic society. This latter step in the analysis is generally itself divided into three additional and concordant analyses, on suitability, on necessity or narrowness, and on strict proportionality.<sup>20</sup> These additional steps are crucial for properly weighing the restrictions and their relationship to the facts justifying them.

It is also perceived that psychic integrity can also be violated in disinformation campaigns, since the dissemination of fraudulent news about traumatic events, crises, threats or catastrophes, especially when non-existent or unfounded, has the power to disproportionately stimulate the increase in general levels of stress, the main element of potentiation and triggering of psychic dysfunctions, such as anxiety disorders, phobias, mental confusion, depression, and burnout syndrome<sup>31</sup>.

Finally, it is also known that disinformation can represent a threat to the moral integrity of the subject, insofar as it can negatively influence people's ethical behavior, either by distorting the individual's self-perception, or by inciting him to practice, before third parties, prejudiced, discriminatory or even violent attitudes against certain ethnic, religious, political or social groups. contributing to the spread of hate speech, intolerance and segregation, thus outraging fundamental ethical and moral values.

One cannot forget the deleterious repercussions of one of the most current products of the massification of hate speech: the so-called digital cancel culture<sup>32</sup> and its implications for the right to personal integrity, because, as has been seen globally, it is a social phenomenon of contemporaneity that intends to remove credibility, affecting prestige, the reputation and dignity of a certain recipient, whether an individual or a legal entity, from attacks and condemnatory judgments in the digital environment, but which often go beyond virtual limits and have immeasurable repercussions in the universe of facts, as attested by examples of people who were physically assaulted as a result of fraudulent news that went viral, such as the lynchings that resulted in the homicides of Mrs. Fabiana Maria (mistaken for a misspoken portrait<sup>33</sup>), in Brazil, and of Mr. Ricardo and Mr. Alberto Flores (mistaken for child kidnappers<sup>34</sup>), in Mexico, among thousands of others.

Disinformation and hate speech have demonstrated substantial power to undermine the protection of honor and the recognition of dignity, provided for in article 11 of the ACHR.

The protection of honor, whether in its objective sense, represented by the understanding of reputation, or in its subjective meaning, associated with the perception of dignity and self-esteem, reveals itself as a simultaneous object of both national protection<sup>35</sup>, sometimes enshrined as a fundamental right, sometimes as a conventional

<sup>31</sup> HARVARD MEDICAL SCHOOL. "Anxiety and Stress Disorders. A guide to managing panic attacks, phobias, PTSD, OCD, social anxiety disorder, and related conditions", Special Health Report, 2021.

<sup>32</sup> An example of digital exile can be illustrated with the suspension of former US President Donald Trump's account on the then "Twitter", now "X", under the justification of "risk of further incitement to violence", according to the platform itself. Available in: [https://blog.x.com/en\\_us](https://blog.x.com/en_us).

<sup>33</sup> Available in: <https://www.unicamp.br/unicamp/ju/670/linchamento-virtual-violencia-real>.

<sup>34</sup> Available in: <https://www.bbc.com/portuguese/salasocial-46206104>.

<sup>35</sup> BRASIL. CONSTITUIÇÃO DA REPÚBLICA FEDERATIVA DO BRASIL DE 1988. Artigo 5º - Todos são iguais perante a lei, sem distinção de qualquer natureza, garantindo-se aos brasileiros e aos estrangeiros residentes no País a inviolabilidade do direito à vida, à liberdade, à igualdade, à segurança e à propriedade, nos termos seguintes:

X – são invioláveis a intimidade, a vida privada, a honra e a imagem das pessoas, assegurado o direito a indenização pelo dano material ou moral decorrente de sua violação;

one, elevated to the *status* of a human right, as illustrated by article 11 of the ACHR, insofar as it states that no one can be the object of arbitrary or abusive interference in their private life, in that of their family, nor of illegal offenses to their honor or reputation, which is why the advance of disinformation and hate speech have been laid bare as a complex instrument for promoting crimes against honor, regulated in Brazil through the Brazilian Penal Code<sup>36</sup>.

The conventional wording, in turn, asserts that no one can be the object of *arbitrary or abusive* interference in his or her private life. Furthermore, the proliferation of fraudulent news may, unjustifiably, violate this protection limit conceived by the conventional law, while the propagation of unhealthy information, such as false, manipulated, modified or untrue news, audios and videos, to the extent that they distort and damage the reputation and dignity of the offended, may cause, Thus, undue, disproportionate and unjustified interference in the legal assets of an individual without any type of legal basis<sup>37</sup>.

In addition, the conventional wording, rightly, is even more concerned with the protection against illegal offenses to its honor and reputation, which reflect aggressions even more serious than the so-called arbitrary or abusive interferences pointed out above, since they enter the universe of illegality, apt to attract, therefore, reprimands not only of a civil nature, but also criminal, such as the dissemination of fraudulent news that intentionally imputes a crime to someone (slander) or that incites hatred and violence (hate speech).

Disinformation and hate speech have demonstrated substantial power to also attack freedom of conscience and religion, provided for in article 12 of the ACHR, since, by manipulating and modifying information, it has contributed to the increase in events of extremism, radicalization and promotion of hate speech<sup>38</sup> against certain religious and ethnic groups, See worldwide massacres<sup>39</sup>, armed conflicts<sup>40</sup> and similar events marked by incitement, crimes and other acts of violence motivated by ethnic-religious tensions.

In this sense, the dissemination of fraudulent news has the power to prevent or hinder the individual formulation of judgments, ideas and opinions either about oneself or about the world that circumscribes it, while it will propagate to the subjects inaccurate, inconsistent informational content that does not attract a judgment of reliability necessary for the use and decision-making. representing, therefore, a relevant obstacle to the exercise of conservation and manifestation of beliefs, especially religious ones.

In a similar sense, it is clear that one of the main negative impacts of the proliferation of hate speech and disinformation campaigns has fallen on the freedoms of thought and expression, provided for in article 13 of the ACHR.

In this regard, it is necessary to distinguish, initially, the boundaries between these two freedoms engraved in the aforementioned precept, because, while freedom of thought identifies the freedom to produce and develop thoughts, beliefs, convictions and opinions, without undue or coercive external interference, and it is, therefore, a right of internal scope of its holder, restricted to his intimate sphere, Like the freedom to have religious, philosophical and political beliefs and preferences, freedom of expression, in turn, comprises the freedom to manifest, expose and verbalize one's thoughts, convictions and beliefs, thus acting in an environment external to the core of its holder, demanding, in this logic, the need for communication and externalization of these wills so that they are known by peers and collectivity.

<sup>36</sup> BRASIL. DECRETO-LEI nº. 2,848, de 7 de dezembro de 1940 – CÓDIGO PENAL. CRIMES CONTRA A HONRA. Artigos 138 a 145.

<sup>37</sup> INTER-AMERICAN COURT OF HUMAN RIGHTS (COURT). *Case Álvarez Ramos vs. Venezuela*. Preliminary Exceptions, Forums, Repairs and Coasts. Judgment of August 30, 2019. Serie C No. 380, P. 101.

<sup>38</sup> UNITED NATIONS. Available in: [https://www.un.org/sites/un2.un.org/files/notohate\\_fact\\_sheets\\_en.pdf](https://www.un.org/sites/un2.un.org/files/notohate_fact_sheets_en.pdf).

<sup>39</sup> Available in: <https://www.amnesty.org/en/latest/news/2022/09/myanmar-facebooks-systems-promoted-violence-against-rohingya-meta-owes-reparations-new-report/>.

<sup>40</sup> For example, the globally known war conflicts between Israel and Palestine, civil war in Syria, Nigeria, India and Pakistan, among others.

From this perspective, despite the fact that freedom of thought is considered an absolute right in many legal systems, including in the inter-American sphere itself<sup>41</sup>, a locus in which thought itself is free and cannot be restricted, the same cannot be said for the verbalization and exteriorization of this thought<sup>42</sup>, that is, freedom of expression, as far as it is concerned, is subject to legitimate limitations, imposed, even, by the conventional wording itself, such as the prohibition of demonstrations in favor of war, as well as any apology for national, racial or religious hatred that constitutes incitement to discrimination, hostility, crime or violence, showing that freedom cannot be used as an instrument of transgression and protection of crimes.

It is a list of restrictions that is consistent with the very core that underpins and grounds the reason for the existence of freedom of expression as a human right, integrated by three parameters: (i) the first, which is firmly based on the premise of seeking "truth and the debate of ideas", a concept currently addressed from the perspective of information integrity, that is, information that translates reality with greater precision, consistency and reliability is more likely to be revealed through the free debate of divergent ideas and opinions – free market of ideas<sup>43</sup> – stripping itself as a focal concept to guarantee the vitality of democratic regimes; (ii) the second, as an expression of the dignity of the human person, since people are only fully realized when they can manifest their convictions and thoughts, an individual dimension of this guarantee; (iii) the third, as the protection and promotion of democracy, a political system in which power is exercised through concrete popular participation and representation, a model that includes not only the right to vote in elections, but also active and effective participation in public debates, demonstrations and other forms of political expression, established in an environment of inclusion of all groups in society, regardless of race, gender, religion, sexual orientation, *socioeconomic status*, as well as the recognition and appreciation of the diversity of opinions, ideologies and interests, ultimately guaranteeing the principle of equality and respect for human rights.

Thus, considering that freedom of expression was designed to serve as a human right that seeks to protect, at the individual level, the dignity and autonomy of the subject and, at the collective level, the functional coexistence in a democratic society<sup>44</sup>, we understand the reasons that require its exercise to be limited so that it does not legitimize conducts that subvert its own foundations, such as the dissemination of fraudulent news that undermines the democratic regime (such as movements to subvert the current Order and attacks against democratic institutions), that promote hate speech and that incite the practice of criminal conduct.

From this perspective, the OAS initially established as follows:

The inter-American human rights system has concluded that freedom of expression is characterized as a right with two dimensions: an individual one, which concerns the expression of one's thoughts, ideas, and information; and a collective or social dimension, consisting of the right of society to procure and receive information, to know the thoughts, ideas, and information of others and to be well informed<sup>45</sup>.

<sup>41</sup> INTER-AMERICAN COURT OF HUMAN RIGHTS (COURT). *Case Ricardo Canese vs. Paraguay*. Forums, Repairs and Coasts. Judgment of August 31, 2004. Series C No. 111, § 77; *Case López Álvarez vs. Honduras*. Forums, Repairs and Coasts. Judgment of February 1, 2006. Serie C No. 141, P. 163.

<sup>42</sup> INTER-AMERICAN COURT OF HUMAN RIGHTS (COURT). *Case Baraona Bray vs. Chile*. Preliminary Exceptions, Forums, Repairs and Coasts. Judgment of November 24, 2022. Serie C No. 481, P. 103.

<sup>43</sup> In this sense, see: STUART MILL, J. *About freedom*, Martin Claret, São Paulo, 2003.

<sup>44</sup> The INTER-AMERICAN COURT OF HUMAN RIGHTS (COURT) has a peaceful position on the centrality of freedom of expression in the existence of a democratic society, in accordance with Advisory Opinion OC 5/85 of November 13, 1985. Available in: [https://www.corteidh.or.cr/docs/opiniones/seriea\\_05\\_por.pdf](https://www.corteidh.or.cr/docs/opiniones/seriea_05_por.pdf).

<sup>45</sup> OAS. "Guide to Guarantee Freedom of Expression in the Face of Deliberate Disinformation in Electoral Contexts, 2019.

In a complementary sense, with the digitalization of the public sphere<sup>46</sup>, the perspective of the right to "seek, receive and disseminate information and ideas of all kinds" changes significantly. The right to information is now thought of not only as access to digital content, but also as access to complete and quality information.

In this vein, in recent years, it is recalled that Brazil has been the scene of serious attacks on the quality of public debate and the weakening of institutional structures and essential public policies, which were thought to be consensual. The COVID-19 pandemic, notably, triggered a process of popularization of the anti-vaccine discourse in Brazil, which came to impact consolidated public policies that were once worthy of broad social trust, such as the National Vaccination Program, without forgetting the attacks on the Praça dos Três Poderes and its respective physical structures, which represented true offenses to the democratic institutions themselves.

In a complementary sense, as will be demonstrated in a separate section that will address the international debate on the governance of digital platforms, it is clear that one of the main current concerns has been the one intended to distinguish freedom of expression from the so-called freedom of *reach*, that is, the right to freedom of expression does not equate to a right to be driven by algorithms. The consolidation of evidence and research consensus by international panels of experts such as the International Panel on the Information Environment has signaled that the digital environments produced by platforms are not neutral and empty spaces, as algorithms actively boost certain discourses to the detriment of others, highlighting and horizontalizing content that was treated as extreme niche discourses, as attested by the discourses that have become more popular in Brazil and in the world in recent years - anti-vaccines and fraud in electoral processes - given that they attack precisely the right to public health and the democratic regime.

Continuing, it should be noted that disinformation can also represent a threat to the right of rectification or reply, enshrined in Article 14 of the ACHR, since this provision provides the recipient of inaccurate or offensive information with the right to correct or respond to it, as a measure of pacification and social justice, thus preventing a reaction measure that may be slow and bureaucratic and could represent an indirect stimulus in an environment that is already very much a propagator of news Fraudulent.

A complex problem of the current digital society has challenged the realization of this conventional guarantee: the stratospheric speed of informational circulation of the inaccurate and offensive message. In a matter of seconds, disinformation summarily, virtually, and socially condemns the recipient, without any kind of minimally critical judgment on the actual authorship or materiality of the potential conduct in disapproval. The punishment then spreads instantly to the *offline world* (plane of facts) causing often irreversible damage, illustrated by the growing culture of *digital cancellation*, *lynching* and *cyberbullying*, materialized by attacks on the means of existence, subsistence and coexistence of the canceled, which usually result in dangerous attacks on their integrity, from dismissals to serious physical aggressions, from suicides or homicides to genocides, emptying the exercise of the right of rectification and reply, since they do not reach with the same effectiveness - temporal and material - the paths and substrates that were initially accessed by the inaccurate or offensive information.

Finally, under the terms of Article 23 of the ACHR, it is recalled that one of the main reflections of disinformation and hate speech disseminated through digital platforms is also that which is laid bare as an instrument of manipulation and intervention in the political process, which aims, in particular, to influence not only the process of electing political representatives, phases prior to and concomitant with the electoral period, as well as the post-election stage of inspection of elective mandates and accountability<sup>47</sup>.

Thus, as demonstrated above, it is the human right of citizens to actively participate in the process of electing their representatives, both through the right to vote and to be voted for, a result that will be guaranteed through equal access to public functions in their

---

<sup>46</sup> COSTA ASSIS, C. "O PARADOXO DA ESFERA PÚBLICA DIGITAL", *Cadernos de Direito Actual*, [S. l.], n. 21, 2023, p. 101-129.

<sup>47</sup> INTER-AMERICAN COURT OF HUMAN RIGHTS (COURT). *Case Moya Chacón y otro v. Costa Rica*. Preliminary Exceptions, Forums, Repairs and Coasts. Judgment of May 23, 2022. C Series No. 451.

country. However, in order for the designated scenario of balance and respect for political rights to be guaranteed, it is essential that citizens are protected from fraudulent information that was intentionally disseminated in order to favor certain groups and candidates to the detriment of others, which is why many countries have shifted their concerns regarding the need to guarantee electoral integrity in the face of disinformation to international environments systematic, seeking some kind of cross-border help for this world-class problem.

At the inter-American level, in October 2019, the OAS published the Guide to Guarantee Freedom of Expression in the Face of Deliberate Disinformation in Electoral Contexts, a document that begins by describing the phenomenon of disinformation in electoral contexts, with an emphasis on some of the benefits brought by this new informational<sup>48</sup> ecosystem, presenting some factors that should be weighed when analyzing the informational disorder produced in the electoral context, such as (i) political polarization (remembering the impact of echo chambers), (ii) the business model of digital platforms anchored in online advertising (disinformation is propagated through the same advertising techniques and tools used by the advertising market) and (iii) viralization (studies show that fraudulent news circulates more quickly than the true news)<sup>49</sup>.

The OAS Guide<sup>50</sup> also discusses the impacts of digital platforms in this context of informational dysfunction, identified as species of the category of intermediaries and platforms:

---

<sup>48</sup> As benefits of the popularization of the Internet mentioned by the OAS in the Guide to guarantee freedom from expression in the face of deliberate disinformation in electoral contexts, the following passages deserve to be highlighted: "La Internet es una red descentralizada que ha expandido la posibilidad de que las personas y los ciudadanos se expresen en la esfera pública, con acceso a recursos tecnológicos mínimos. Este nuevo ecosistema de comunicación presenta enormes oportunidades para la expansión y el fortalecimiento del debate público, así como para la circulación y acceso a la información de interés público (...) Internet es una de las tecnologías que más ha potenciado el ejercicio de la libertad de expresión, dado que convirtió a millones de personas que eran receptores pasivos de información en activos participantes del debate público."

<sup>49</sup> (...) Polarization. Several studies link the phenomenon of disinformation to political polarization. Some argue that disinformation is the cause of polarization, as it appeals to people's emotions and seeks to affect their behavior, including voting behavior (such as so-called "voter suppression," i.e., practices or strategies used to discourage certain sectors of the population from participating in the electoral process)<sup>7</sup>. Others argue that disinformation is actually the effect of political polarization, as people lock themselves into narrow information silos in which they only access information through the prism of their own ideologies or political positions (spaces called "echo chambers"), which prevents them from accessing divergent points of view (...) Thus, it has been argued that the online advertising market, based on the mining of personal data and the ability to target messages accurately and effectively, encourages disinformation as it is spread through the same tools and techniques used by the advertising market.<sup>15</sup> In this regard, the use of personal data for advertising purposes—largely permitted by outdated or limited data protection laws—appears to create conditions that favor the phenomenon or make it more effective. It should be noted here that many of these norms in the region could be outdated in relation to these new phenomena or their application could be limited. Internet virality. It should be noted that the phenomenon of Internet virality is not in itself problematic: on the contrary, by taking advantage of the decentralized nature of the network, the circulation of information can bypass controls or points of blockage or censorship, something especially valuable in the contexts of authoritarian countries. However, when it comes to misinformation, it can be a factor that complicates the approach to the phenomenon. There are studies that suggest that fake news circulates faster than the real news, because they appeal to emotions and are - consequently - more attractive.<sup>18</sup> It has also been argued that social networks offer a form of information exchange that is "epistemically valuable": a publication on a social network is equivalent to a "testimony" given by another person, usually someone we know or appreciate for some reason (that is why we "follow" or are his "friends", etc.). As human beings tend to accept what others transmit to us in the form of testimony, it is possible to postulate that disinformation flourishes in social networks because it takes advantage of an epistemologically valid procedure for most cases and uses. OAS. "Guide to Guarantee Freedom of Expression in the Face of Deliberate Disinformation in Electoral Contexts".

<sup>50</sup> *Idem*, p. 21.

The paradigm shift in the information flow of democratic societies was accompanied by the emergence of new actors. That is the case of intermediaries and platforms operating on the Internet seem to play central roles in the new information ecosystem. These actors have had a positive impact on their development, by facilitating access to information that we do not know (in the case of search engines), offering the necessary infrastructure to participate in the public debate (such as Internet service providers) or offer platforms where it is possible to share information, ideas, and access content produced by third parties (as is the case with media consumption platforms or social networks). These services, managed by the private sector, have become important actors, partly as a result of the concentration - in market terms - prevalent in these services.<sup>2</sup>

The legacies of intermediaries and platforms in the construction of a more plural debate space have not only been positive, as there are known and challenging risks involved in this mission, especially when faced with attempts at arbitrary and unjustified control of digital discourse, either through censorship or through manipulation – processing that compromises the integrity – of electoral information. Instruments now commonly recruited by authoritarian leaders and governments in order to forge the flow of information to serve their own interests, euphemistically as foreseen by the OAS Guide:

Through their role as intermediaries in the flow of information on the Internet, these actors operate as traffic "control nodes." Their decisions have a significant impact on the public debate, and so a number of governments—especially authoritarian ones—have tried to pressure them to help shape the flow of information according to their interests<sup>51</sup>

Still from the perspective of the electoral process, it should be noted that, in 2020, the UN and the OAS issued a joint declaration, entitled the Joint Declaration on Freedom of Expression and Elections in the Digital Age<sup>52</sup>, in which they highlighted the importance of freedom of expression and information for the electoral system:

(...) deliberate disinformation, misinformation or manipulation, disseminated with the intention of harm, as well as 'hate speech', can generate and aggravate tensions related to elections; calling on parties and candidates to avoid intentionally using such statements to boost their electoral prospects, and recognising the important role that independent electoral authorities and bodies play in addressing these forms of expression and promoting access to information;

Furthermore, examples of manipulation of videos, images and sounds of candidates (deep fakes<sup>53</sup>), unduly associating them with reprehensible and perhaps criminal conduct, continue and will continue to circulate intentionally and virally on social networks, compromising, in addition to the suitability of the electoral process, the integrity of the democratic regime itself.

In this vein, the 2023 report of the International Panel on the Information Environment (IPIE), produced as a result of a survey carried out with experts on the subject

---

<sup>51</sup> *Ibidem*.

<sup>52</sup> Statement issued by the United Nations (UN) Special Rapporteur on Freedom of Opinion and Expression, the Representative on Freedom of the Media of the Organization for Security and Co-operation in Europe, and the Special Rapporteur for Freedom of Expression of the Organization of American States (OAS). ORGANIZATION OF AMERICAN STATES. "JOINT DECLARATION ON FREEDOM OF EXPRESSION AND ELECTIONS IN THE DIGITAL AGE", 2020. Available in: [https://www.ohchr.org/sites/default/files/Documents/Issues/Opinion/JointDeclarationDigitalAge\\_30April2020\\_EN.pdf](https://www.ohchr.org/sites/default/files/Documents/Issues/Opinion/JointDeclarationDigitalAge_30April2020_EN.pdf).

<sup>53</sup> Cf.: SHAO, G. "What deepfakes are and how They may be dangerous", CNBC, 2019; KNIGHT, Will. "The US military is funding an effort to catch deepfakes and Other AI trickery", *MIT Technology Review*, 2018.

around the globe, highlights electoral disinformation as one of the main forms of systemic disinformation, second only to health disinformation<sup>54</sup>.

The proliferation of theses on electoral fraud in digital environments and their offline reverberations, such as the violent contestation of the electoral result on January 8, 2023 in Brazil, focus on the right of citizens to freely choose their representatives.

In the same sense, the agenda of political violence against women candidates or human rights defenders has also been well documented<sup>55</sup>, with digital media being the main environment for the suppression of these rights. In particular, it is noted that hate speech is crossed by other intersectionalities in addition to misogyny, such as racism, transphobia, lgbtphobia, among others. The IACHR has adopted precautionary measures on the subject<sup>56</sup>, a signal that has been directing the internal debate on the subject, in view of the reinforced obligation imposed on the State to protect human rights defenders.

Ultimately, the problems posed by the lack of regulation of the industry go beyond elections, encompassing threats to the sovereignty of States and their operational capacity to enforce the rights of their citizens and other people in their territory. In the absence of regulation, the digital environments produced by the platforms become spaces of parallel sovereignty, where private contractual rules of technology companies that operate in a separate from the application of national laws, including those related to the exercise of human rights, are in force. In this regard, the same OAS Guide<sup>57</sup> highlights as follows:

It has also been found that Internet companies themselves moderate the content that circulates on their platforms, adopt decisions to suppress or reduce the circulation of content according to their own terms of service and "community" rules and thus also form the public debate. Actualmente, y fruto de desafíos como el de la desinformación intencionada y el discurso que ininci al odio, muchas de estas empresas están travesando procesos de revisión de los criterios que utilizan para moderar contenidos, los procedimientos internos que utilizan a tal fin y la selección de información que generan y presentan a los usuarios, basadas en la utilización de algoritmos. This dynamic of public prisons and private actions is aggravated by concentration: as powerful as the actor who operates as an intermediary, more attractive is the actor who operates as an intermediary, more attractive is vuelve for those who want to press and have more impact on their decisions.

Moving forward, it should be noted that disinformation can also represent a risk to the right to equality before the law, engraved in article 24 of the ACHR, which is responsible for providing that everyone has the right, without discrimination, to equal protection of the law. In this vein, the concepts of echo chamber and bubble filters receive prominent attention, as they reveal how certain algorithmic arrangements are designed by platforms and how they will have repercussions beyond digital borders.

The term echo chamber, derived from the English expression echo chambers, identifies, based on an analogy to the sounds that reverberate in a hollow envelope, such

54 INTERNATIONAL PANEL ON THE INFORMATION ENVIRONMENT. "Trends in the Global Information Environment 2023 Expert Survey Results", Synthesis Report 2023.2, 2023.

55 LAURIS, E.; HASHIZUME, M. *Violência Política e Eleitoral no Brasil: panorama das violações de direitos humanos de 2016 a 2020*, Terra de Direitos e Justiça Global, Curitiba, 129 p. Available in: [http://www.global.org.br/wp-content/uploads/2020/09/Relat%C3%B3rio\\_Violencia-Politica\\_FN.pdf](http://www.global.org.br/wp-content/uploads/2020/09/Relat%C3%B3rio_Violencia-Politica_FN.pdf). See also: ASSIS, M. P. et al. "E ainda nos levantamos: luta e resistência das mulheres em defesa dos direitos humanos". In: *Comitê Brasileiro de Defensoras e Defensores de Direitos Humanos. Dossiê Vidas em Luta: criminalização e violência contra defensoras e defensores de direitos humanos no Brasil*, v. III. Organização: Layza Queiroz Santos et al, Terra de Direitos, Curitiba, 2020. Available in: <https://terradedireitos.org.br/uploads/arquivos/Dossie-Vidas-em-Luta.pdf>.

56 IACHR. Resolution No. 34/2022. Precautionary Measure No. 408-22. Benny Briolly Rosa da Silva Santos and members of his work team in relation to Brazil, July 11, 2022. Available in: [https://www.oas.org/pt/cidh/decisiones/mc/2022/res\\_34-22%20\\_mc\\_408-22\\_br\\_pt.pdf](https://www.oas.org/pt/cidh/decisiones/mc/2022/res_34-22%20_mc_408-22_br_pt.pdf).

57 ORGANIZATION OF AMERICAN STATES (OAS). "Guide to Guarantee Freedom of Expression in the Face of Deliberate Disinformation in Electoral Contexts", p. 22.



as bells, a linked and closed media space that has the potential to amplify the messages delivered there and isolate them from the messages that contradict them (avoids dissent), which is why it is not uncommon to confuse it with the expression bubble filters), a phrase that refers to the personalization of results from search engines, such as Google and social media feeds, which are creating a unique universe of information for each of us based on our tastes (profiling<sup>58</sup>).

In addition to the structural network dimension of the platforms mentioned above, it is also worth remembering the practical effects supported by the right to equality as a result of the lack of transparency in the application of content moderation rules by the platforms, a gap that for a long time was anchored in the myth of the supposed exemption and neutrality of the networks (which would be mere intermediaries), but in the face of the disclosure of recent scandals<sup>59</sup>, especially after 2016, they weakened this thesis. In summary, such events showed that the discourse on the platforms results from a complex interaction between users' interests, platform resources and generative algorithmic choices<sup>60</sup>, which exclusively serve the final pursuit of increasing the time users spend on social networks, by capturing their attention.

In this way, there is a content curation produced by the networks' own algorithms that, in interaction with the behavior of users, defines which types of posts will be amplified and which will be restricted (logic of the echo chambers above), having as a measure the engagement capacity of a given content (privileging mainly sensationalist and false content), which is why it would not be minimally reasonable to impute to social networks, in isolation, the duty to mitigate the negative externalities caused by a model of economic exploitation that depends on this logic to survive, to the extent that the eventual effectiveness of the self-regulatory policy of moderation will always coexist with the shadow of conflict of interest (moral hazard).

Furthermore, the absence of clear and transparent rules adopted by social networks in the elaboration and, especially, application of their content moderation policies promotes, therefore, an environment of unjustified differential treatment for certain groups of users, since individuals with greater capacity for engagement have greater permissiveness and leniency from companies to publish fraudulent news or news that incites violence, for example, when compared to those with lower levels of visibility, an asymmetry that fosters, therefore, a legally discriminatory informational ecosystem.

Designated asymmetries call for coordinated state and regional responses, especially through the promotion of promotion measures that intend to raise the levels of literacy and digital literacy<sup>61</sup>, as they will have the potential to increase the levels of collective awareness of the population about the benefits and risks of digital platforms, an objective usually materialized through the implementation of public policies, because, as observed throughout the present research, the higher the level of organic and social maturity of a given State, the lower the pernicious impacts of the informational disorder, in addition to the lower possibility that the (intentional) human circulation will propagate fraudulent news or legally abusive speeches - users will enjoy greater critical judgment and knowledge about the risks involved.

In a complementary sense, with regard to the Judiciary of the OAS member states, in turn, it has been found that the absence of a clear regulatory normative parameter has conferred on them a double function, that is, the first, as a space that is increasingly called

---

<sup>58</sup> DONEDA, D. *Da privacidade à proteção de dados pessoais*, Renovar, Rio de Janeiro, 2006, p. 173.

<sup>59</sup> MIT TECHNOLOGY REVIEW. "How Facebook got addicted to spreading misinformation", 2021. Available in: <https://www.technologyreview.com/2021/03/11/1020600/facebook-responsible-ai-misinformation/>.

<sup>60</sup> DOUEK, E. "Governing Online Speech: From 'Posts-as-Trumps' to Proportionality and Probability", *Columbia Law Review*, [s.1], v. 121, n. 759, 2020.

<sup>61</sup> Terms that encompass a wide range of cognitive, social, motor, and emotional skills that users need to enjoy in order to navigate the digital environment effectively, such as access to digital education policies that effectively transfer knowledge, educate, and raise awareness among users so that they can identify what is disinformation, malicious information, among other data processing that compromises the integrity of information. It is also required to deepen the channels available for reporting illegal, hateful or violent content as a mechanism of social control.

upon to intervene and pacify these controversies, and the second, as a reflection of this broad protagonism, that ends up requiring the national Judiciary to impose conditions - isolated<sup>62</sup>, however - on the moderation policies of social media platforms, creating a particularly problematic and casuistic conjuncture if one considers that only a tiny fraction of the debates involving the limits to the exercise of freedom of expression on social networks are judicialized (problem of selectivity), making it impossible to have a real and systematic confrontation of the problem that is proportional to the dimension of the unprecedented political and market power of these giant economic agents that compete with the State and can certainly influence and affect the public interest.

Finally, firmly in this context, there is growing social pressure for the Legislative Powers of the OAS member states to claim, evoke and execute their primary functions as a natural locus of democratic representation of the interests of society and the construction of the legislative framework necessary to convey the rules that will inform this new niche of activities, scenarios that seem to converge to the creation of this whole of rules and principles. harmonized nationally and regionally, to illuminate the governance of digital platforms in full harmony with the dictates of the Inter-American Human Rights System.

#### 4 CONCLUSIONS

As exposed, with the advent and consolidation of the Informational Order, the relationship of digital subjects with information was greatly impacted. The reasons for this significant change have several perspectives, however, *in casu*, those that identify the association of the phenomena of globalization and the technological revolution deserve prominence, regardless of the terminologies used to describe these events (Fourth Industrial Revolution, Surveillance Capitalism, digitalization of private life and virtualization of intimacy), since it is certain that the Information Society has produced intense changes not only social and political, but, above all, in the way human beings relate to law and space.

In fact, it was inferred that the main threats to information security, especially disinformation and hate speech perpetrated through digital platforms, emerge as products of this new Informational Order, attracting, in addition to the well-known benefits of digital transformation, transnational risks, revealing that digital platforms operate beyond national borders, with actions that have global repercussions. This cross-border trait of disinformation requires enhanced international cooperation and harmonization of legal treatment so that the challenges are addressed effectively, thus ensuring that the digital human rights provided for in the ACHR are effectively protected.

The research showed that, although there are regulatory frameworks and initiatives aimed at mitigating the unwanted effects of disinformation, there are still significant gaps in the application and effectiveness of these measures. The insufficiency of normative arrangements that discipline transparent rules on issues that are core to the governance of platforms, such as content moderation and accountability for their conduct, associated with the lack of integrated public policies for digital literacy and collective awareness, end up causing uncoordinated and insufficient jurisdictional responses, which is why these vulnerabilities have been identified as obstacles to the effective protection of human rights Digital.

In view of this scenario, the present research reaffirms the need for transnational, systematic and multidimensional solutions to face the challenges imposed by digital disinformation, since, based on what is evidenced, only through an integrated approach, which brings together effective governance, education and international cooperation, will it be possible to promote a more functional informational ecosystem, strengthening

---

<sup>62</sup> Considering the perspective of being adopted through random and disconnected decision-making expedients, coming from the most different levels of the Brazilian Judiciary, see examples of decisions rendered both by lower court judges who prevented the operation of Whatsapp or Telegram, as well as the most recent and famous commands issued by the Brazilian Federal Supreme Court against the former *Twitter*, today *X*.

democracy and ensuring the protection and promotion of rights provided for in the Inter-American System for the Protection of Human Rights (IAHRS).

## REFERENCES

- BELL, D. *The Coming of Post-Industrial Society: A Venture in Social Forecasting*, Basic Books, New York, 1973.
- BRASIL. CONSTITUIÇÃO DA REPÚBLICA FEDERATIVA DO BRASIL DE 1988.
- BRASIL. DECRETO-LEI Nº 2.848, DE 7 DE DEZEMBRO DE 1940. CÓDIGO PENAL.
- BRASIL. RIO GRANDE DO SUL. DECRETO Nº 57.603, DE 5 DE MAIO DE 2024.
- CARVALHO, M. S. R. M. *A trajetória da Internet no Brasil: do surgimento das redes de computadores à instituição dos mecanismos de governança*. Dissertação de mestrado, Programa de Pós-Graduação em Engenharia de Sistemas de Computação, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2006.
- CASTELLS, M. *A sociedade em rede (A era da informação, economia, sociedade e cultura)*, Paz e Terra, São Paulo, 1999.
- CLEMENTINO, M. B. M. *A cooperação jurídica internacional em matéria penal-tributária como instrumento de repressão à criminalidade organizada transnacional: globalização e novos espaços de juridicidade*. Tese de Doutorado em Direito, Programa de Pós-Graduação em Direito, Universidade Federal de Pernambuco, Recife, 2013.
- INTER-AMERICAN COMMISSION ON HUMAN RIGHTS (IACHR). *Desinformación, pandemia y derechos humanos*, 2023.
- INTER-AMERICAN COMMISSION ON HUMAN RIGHTS (IACHR). Resolution No. 34/2022. Precautionary Measure No. 408-22. Benny Briolly Rosa da Silva Santos and members of his work team in relation to Brazil, July 11, 2022.
- INTER-AMERICAN COURT OF HUMAN RIGHTS (COURT). *Case Álvarez Ramos vs. Venezuela*. Preliminary Exceptions, Forums, Repairs and Coasts. Judgment of August 30, 2019. Serie C No. 380., P. 101.
- INTER-AMERICAN COURT OF HUMAN RIGHTS (COURT). *Case Baraona Bray vs. Chile*. Preliminary Exceptions, Forums, Repairs and Coasts. Judgment of November 24, 2022. Serie C No. 481., P. 103.
- INTER-AMERICAN COURT OF HUMAN RIGHTS (COURT). *Case Moya Chacón y otro v. Costa Rica*. Preliminary Exceptions, Forums, Repairs and Coasts. Judgment of May 23, 2022. C Series No. 451.
- INTER-AMERICAN COURT OF HUMAN RIGHTS (COURT). *Case Ricardo Canese vs. Paraguay*. Forums, Repairs and Coasts. Judgment of August 31, 2004. Series C No. 111, P. 77.
- INTER-AMERICAN COURT OF HUMAN RIGHTS (COURT). *Case López Álvarez vs. Honduras*. Forums, Repairs and Coasts. Judgment of February 1, 2006. Serie C No. 141, P. 163.
- INTER-AMERICAN COURT OF HUMAN RIGHTS (COURT). Advisory Opinions OC 5/85, of November 13, 1985. Available in: [https://www.corteidh.or.cr/docs/opiniones/seriea\\_05\\_por.pdf](https://www.corteidh.or.cr/docs/opiniones/seriea_05_por.pdf).
- COSTA ASSIS, C. "O PARADOXO DA ESFERA PÚBLICA DIGITAL", *Cadernos de Direito Actual*, [S. l.], n. 21, 2023, p. 101–129.
- DIVINO, S. "Tutela Ética e Jurídica dos Danos e da Responsabilidade da Inteligência Artificial", *Cadernos de Direito Actual*, [S. l.], n. 22, p. 288–313, 2023.
- DONEDA, D. *Da privacidade à proteção de dados pessoais*, Renovar, Rio de Janeiro, 2006.
- DOUEK, E. "Governing Online Speech: From 'Posts-as-Trumps' to Proportionality and Probability", *Columbia Law Review*, [s.l.], v. 121, n. 759, 2020.
- EUROPEAN COMMISSION, Directorate-General for Communications Networks, Content and Technology, "A multi-dimensional approach to disinformation – Report of the independent High level Group on fake news and online disinformation, Publications Office", 2018, Available in: <https://data.europa.eu/doi/10.2759/739290>.

- GANDOMI, A.; MURTAZA, H. "Beyond the hype: Big data concepts, methods, and analytics", *International Journal of Information Management*, v. 35, n. 2, 2015, p.137-144, ISSN 0268-4012, <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>.
- HARVARD MEDICAL SCHOOL. "Anxiety and Stress Disorders. A guide to managing panic attacks, phobias, PTSD, OCD, social anxiety disorder, and related conditions", Special Health Report, 2021.
- HOOFNAGLE, C. J.; SOLTANI, A.; GOOD, N.; WANBACH, D. J. "Behavioral advertising: the offer you can't refuse", *Harvard Law & Policy Review*, v.6, 2012, p. 273-296.
- INTERNATIONAL PANEL ON THE INFORMATION ENVIRONMENT. "Trends in the Global Information Environment 2023 Expert Survey Results", Synthesis Report 2023.2, 2023.
- INTRONA, L.; NISSEBAM, H. "Shaping the web: why the politics of search engines matters", *The Information Society*, v. 16, 2000.
- KNIGHT, Will. "The US military is funding an effort to catch deepfakes and Other AI trickery", *MIT Technology Review*, 2018.
- LAURIS, E.; HASHIZUME, M. *Violência Política e Eleitoral no Brasil: panorama das violações de direitos humanos de 2016 a 2020*, Terra de Direitos e Justiça Global, Curitiba, 2020.
- MIT TECHNOLOGY REVIEW. "How Facebook got addicted to spreading misinformation", 2021.
- ORGANIZATION OF AMERICAN STATES (OAS). "Guide to Guarantee Freedom of Expression in the Face of Deliberate Disinformation in Electoral Contexts", 2019.
- ORGANIZATION OF AMERICAN STATES (OAS). "JOINT DECLARATION ON FREEDOM OF EXPRESSION AND ELECTIONS IN THE DIGITAL AGE", 2020.
- OLIVEIRA, R.R.V; CLEMENTINO, M.B.M. "Os MLATS vigentes no Brasil e as implicações na soberania contemporânea: internacionalização, territorialidade e jurisdição", *Revista FIDES*, v. 12, n. 1, 2021, p. 830-849.
- OLIVEIRA, R. R. V. *Regime brasileiro de transferência internacional de dados pessoais*, Lumen Juris, Rio de Janeiro, 2024
- PAN AMERICAN HEALTH ORGANIZATION. "Understanding the Infodemic and Misinformation in the fight against COVID-19". (<https://iris.paho.org/handle/10665.2/52052>).
- PASQUALE, F. *The black box society. The secret algorithms that control money and information*, Harvard University Press, Cambridge, 2015.
- SCHWAB, K. *The Fourth Industrial Revolution*, Penguin Books, London, 2017.
- SCHWAB, K.; DAVIS, N. *Shaping the Fourth Industrial Revolution*, World Economic Forum, Geneva, 2018.
- SHAO, G. "What deepfakes are and how They may be dangerous", CNBC, 2019.
- SKAFLE, I., NORDAHL-HANSEN, A., QUINTANA, D. S., WYNN, R., & GABARRON, E. "Misinformation About COVID-19 Vaccines on Social Media: Rapid Review", *Journal of Medical Internet Research*, 2022.
- STUART MILL, J. *Sobre a liberdade*, Martin Claret, São Paulo, 2003.
- UNITED NATIONS. "Policy Brief for Our Common Agenda: Information Integrity of Digital Platforms", 2023.
- UNITED NATIONS. "Strategy and plan of action on hate speech", 2019.
- VERONESE, A. "The fourth industrial revolution and blockchain: social values and trust", *ADENAUER NOTEBOOKS XXI*, Konrad Adenauer Foundation, Rio de Janeiro, 2020.
- VON BOGDANDY, A. "Globalization and Europe: How to Square Democracy, Globalization, and International Law", *European Journal of International Law*, v. 15, n. 5, 2004.
- VOSOUGHI, S.; ROY, D; ARAL, Sinan. "The Spread of True and False News Online", 2018. Available in: <https://science.sciencemag.org/content/359/6380/1146>.
- WORLD HEALTH ORGANIZATION. "Disinformation and public health". Available in: <https://www.who.int/news-room/questions-and-answers/item/disinformation-and-public-health>.
- ZUBOFF, S. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, Profile Books, London, 2019.