UDC: 323.266 DOI: 10.56461/iup_rlrc.2024.5.ch6

Fernanda F. FERNANDEZ JANKOV*[*]
Law Faculty, University of São Paulo, Brazil
Fernandez & Jankov Legal Intelligence

TOWARDS A GLOBAL REGULATORY REGIME FOR TECH GIANTS

This paper offers tools for regulatory authorities to effectively address the spread of fake news by tech giants. Evaluating current frameworks, which often focus on symptom treatment like content removal and fact-checking, the study finds these methods insufficient for tackling the root causes of misinformation. Proposing a harm-based regulatory regime inspired by social medicine, political science, and legal theory, the paper emphasizes a holistic approach. Integrating insights from political science and revisiting the concept of regimes as global regulation, it provides a structured framework for regulatory authorities. This approach includes understanding socio-economic incentives, leveraging advanced technologies like AI, and promoting digital literacy. The study highlights the importance of principles, norms, rules, and decision-making processes to create a coherent regulatory environment adaptable to various socio-political contexts where interdisciplinary collaboration among governments, digital platforms, civil society, and international organizations is crucial. The proposed regime aims to foster a trustworthy information ecosystem, enhance societal trust, and mitigate the impact of fake news. By recognizing the complexity of fake news, this paper provides mechanisms for raising awareness among all actors involved and structuring their actions within such a legal framework. The ultimate aim is to establish a resilient and reliable digital public sphere, offering regulatory authorities a comprehensive strategy to combat digital misinformation effectively.

Keywords: tech giants regulatory regime, digital literacy, fake news, harm-based approach, trustworthy information ecosystem.

[*] PhD, post-doctorate programme, ORCID: 0009-0007-9737-4293, e-mail: *fernanda.jankov@usp.br*

1. INTRODUCTION

This paper offers tools for regulatory authorities to effectively address the spread of fake news by tech giants, providing insights on how to legally structure a regulatory regime using ideas from political science and legal theory.[1] In the digital age, the proliferation of fake news poses a significant challenge to the integrity of information disseminated online, particularly by tech giants. This paper examines the regulatory mechanisms aimed at holding these corporations accountable for the spread of misinformation, evaluating their effectiveness as deterrents. Current frameworks often focus on symptom treatment, such as content removal and fact-checking, rather than addressing the root causes and systemic issues that allow misinformation to thrive.

Drawing parallels between medical diagnosis and legal regulation, this study argues for a holistic and integrated approach to combating fake news. By incorporating insights from social medicine, political science, and legal theory, we propose a harm-based regulatory regime that addresses the multifaceted nature of the digital media landscape. This approach emphasizes the importance of understanding the socio-economic incentives and psychological factors that drive the production and dissemination of fake news, as well as the role of digital platforms in amplifying misleading content.

The paper advocates for the development of a comprehensive strategy that includes robust legal frameworks, educational initiatives, and technological solutions. Key to this strategy is the promotion of digital literacy programs to equip users with the skills needed to critically assess information. Additionally, leveraging advanced technologies such as artificial intelligence and machine learning can help detect and mitigate the spread of fake news.

By fostering a collaborative environment involving governments, private digital platforms, civil society, and international organizations, this paper aims to create a more resilient and trustworthy information ecosystem. The proposed regulatory regime emphasizes the importance of principles, norms, rules, and decision-making processes to create an environment for a coherent regime that is adaptable to various socio-political contexts. This interdisciplinary collaboration ensures that the regulatory measures are context-sensitive and effective.

Ultimately, the proposed regulatory regime seeks not only to hold tech giants accountable but also to restore societal trust in information and enhance the overall resilience of digital information ecosystems. By recognizing the complexity of fake news, this paper provides mechanisms for raising awareness among all actors involved and structuring their actions within such a legal framework. This study contributes to the ongoing discourse on effective legal strategies for combating digital misinformation, aiming to establish a resilient and reliable digital public sphere.

[1] Disclaimer: This paper does not aim to offer a complete Regulatory Regime for Tech Giants but rather to provide the foundational basis and legal structure from the perspective of legal thinking. It is intended to serve as a seed for debate and to enhance understanding of the complexity involved in addressing the regulation of Tech Giants. The ideas presented herein lay the groundwork for a broader research project that involves comparative law systems and interdisciplinary approaches. The purpose of this publication is to stimulate scholarly discussion and contribute to the ongoing discourse on this critical issue.

2. UNDERSTANDING FAKE NEWS AND ITS IMPACT

2.1. Defining Fake News

Fake news refers to misinformation or disinformation that is intentionally spread to deceive the public. According to Abiri & Buchheim (2022), fake news is not merely false information but a deliberate distortion intended to manipulate public perception. This understanding highlights the intentional aspect of fake news, distinguishing it from mere errors or inaccuracies in reporting. It is this deliberate intent to mislead that separates fake news from other forms of incorrect information.

Such manipulation is facilitated by the digital environment, which allows for rapid dissemination and amplification of misleading content. Moreover, as Abiri & Buchheim (2022) point out fake news exploits the digital epistemic divide, where different segments of the population have varying access to and trust in information sources. This divide is exacerbated by algorithmic filtering, which creates echo chambers and reinforces pre-existing beliefs. This scenario contributes to the challenge of distinguishing between credible and non-credible sources, further complicating the fight against fake news.

The aspect of what it serves is addressed in the definition offered by Humprecht (2018, p. 3) where fake news refers to "online publications of intentionally or knowingly false statements of facts that are produced to serve strategical purposes and are disseminated for social influence or profit." In this sense, the examination of the characteristics of fake news leads to the assertion that it is often produced and disseminated for strategic purposes, either ideological or commercial aiming to change recipients' perceptions of certain issues and, in the long run, influence their opinions or behaviour.

Humprecht (2018, p. 3) categorizes fake news into several types, including satire, parody, fabrication, manipulation, propaganda, and rumours or hoaxes. Satire and parody involve humour or exaggeration to critique or mock real events, which can sometimes be mistaken for factual news. Fabrication refers to completely false information created to deceive readers, while manipulation involves distorting or altering facts to mislead. Propaganda is biased or misleading information used to promote a political cause or point of view. Rumours and hoaxes are unverified pieces of information that spread rapidly, often causing public alarm or outrage. Based on Calil (2022) expands on these categories by discussing the role of public agents in social media regulation, arguably highlighting that fake news can also include misleading political statements and false narratives spread by political actors to manipulate public perception.

Approaching fake news from this perspective reveals that the issue extends beyond a simple dichotomy of true versus false information. Fake news is intricately linked to the concept of legitimacy, as it undermines the authority and credibility of "central" institutions, both political and scientific. This paper introduces the term "fake legitimacy" to describe the type of legitimacy that arises from such manipulation. Unlike genuine legitimacy, which is grounded in truth and authenticity,[2] 'fake legitimacy' stems from

[2] In the pre-digital era, information dissemination was primarily based on a broadcasting model, where information was distributed from a single source to a broad audience. This model was subject to higher levels of accountability as broadcasters were regulated by stringent legal and ethical standards, ensuring

misinformation and deception, creating a perception that does not align with reality. The creation of this "fake legitimacy"[3] prevents the establishment of a trustworthy information ecosystem and further erodes societal trust, which is essential for democratic governance.

2.2. Why Fake News Exists as a Disease and Its Symptoms

In examining fake news through the lens of a disease and its symptoms, this study aims to leverage interdisciplinary insights, particularly those proposed by Stephenson & Rinceanu (2023). Their work explores the historical and ongoing synergy between law and medicine, advocating for an integrated approach to internet regulation. Drawing on the ideas of legal realists like Oliver Wendell Holmes Jr. and Benjamin Cardozo, they argue that effective solutions to global internet regulation require the combined efforts of medical and legal professionals to address online social problems. This interdisciplinary approach is essential for understanding the epistemic changes brought about by digital media and for developing effective regulatory frameworks. The European Union's "notice-and-takedown" model and North America's "market self-regulation" model, for instance, represent different approaches to regulating online communications, highlighting the profound disagreements on free speech's role in democratic governance (Stephenson & Rinceanu, 2023).

By conceptualizing fake news as a disease, this paper emphasizes the need to diagnose and address the root causes of the digital epistemic divide rather than merely treating the symptoms of misinformation. This involves adopting sophisticated harm-based approaches that rebuild trust in epistemic institutions, integrate free speech theories, and leverage interdisciplinary insights from both law and medicine to create effective regulatory frameworks. By fostering transparency, accountability, and inclusive dialogue, these solutions aim to restore the common factual ground necessary for democratic legitimacy and social cooperation. This holistic approach not only addresses the immediate impacts of fake news but also seeks to understand and mitigate the underlying conditions that allow such misinformation to proliferate.

The drivers of fake news production and diffusion are multifaceted. At the individual level, according to Humprecht (2018, p. 3), psychological effects such as confirmation bias and motivated reasoning lead people to believe information that confirms their existing beliefs. This new environment is marked by a shift from an offline ‘broadcasting' to an online ‘participatory' communication model and, second, the rise of dominant, privately owned digital intermediaries, the so-called ‘Big Five’ (namely, Alphabet/formerly Google, Meta/formerly Facebook, Microsoft, Amazon, Apple).[4] Humprecht' s

the accuracy and reliability of the information provided. The shift to a participatory model in the digital age, where anyone can create and share content, has significantly reduced these accountability mechanisms, facilitating the spread of misinformation and the rise of 'fake legitimacy.'

[3] The term 'fake legitimacy' has been introduced in this paper as part of a broader research project. This project aims to further explore how the concept of legitimacy has evolved in the digital era, particularly in the context of misinformation and the influence of digital platforms.

[4] Max Planck Institute for the Study of Crime, Security and Law. 2024. *Rethinking Digital Media Regulation*. Available at: https://csl.mpg.de/en/projects/rethinking-digital-media-regulation?c=178896 (27. 6. 2024).

(2018, p. 3) research highlights that social media is often used for its entertainment value, which contributes to the uncritical dissemination of misleading information. Moreover, people tend to trust information from sources that align with their pre-existing beliefs, further fuelling the spread of fake news.

At the societal level, still follow the same study. Humprecht (2018, p. 10), factors such as media trust, political polarization, and the strength of public service broadcasting (PSB) significantly influence the prevalence and impact of fake news leading to the ascertainment that countries with strong PSB and higher levels of trust in government and professional news media tend to have lower levels of partisan disinformation. Conversely, countries with lower media trust and higher political polarization, such as the United States and the United Kingdom, experience higher levels of partisan fake news. This relationship underscores the role of institutional trust and the media environment in either mitigating or exacerbating the spread of fake news.

Another contributing factor is the media ecosystem as Abiri & Buchheim (2022, p. 45) add. A fragmented media landscape with varying journalistic standards enables the proliferation of fake news. Social media platforms, in particular, play a crucial role in spreading false information. These platforms often lack the rigorous editorial oversight found in traditional news organizations, allowing misinformation to circulate widely and quickly. The ease with which content can be shared and the algorithms that prioritize engaging (often sensational) content further exacerbates the issue (Caled & Silva, 2022).

Economic incentives also drive the spread of fake news. Fake news can be economically profitable, as sensational stories attract clicks and generate ad revenue. This financial motivation incentivizes the creation and dissemination of false content. Entities that produce fake news often prioritize virality over accuracy, exploiting the economic benefits of high engagement rates (Stephenson & Rinceanu, 2023, p. 79).

Political polarization, a fragmented media ecosystem, and economic incentives, collectively contribute to the resilience and spread of fake news. By understanding these underlying causes, it becomes evident why fake news exists as a persistent "disease" in the information landscape, creating an environment where misinformation can thrive, and challenging efforts to maintain an informed and cohesive society. This understanding is essential for this study as it aims at setting the core basis of a Regulatory Regime of Tech Giants which must address these specific issues if it is to be effective.

As for the symptoms of fake news, they manifest in various detrimental ways, extending far beyond the mere dissemination of false information as it erodes trust in traditional media and democratic institutions, contributes to societal polarization, and undermines the shared factual basis necessary for effective public discourse and policymaking, essential elements of legitimacy as relevant studies demonstrate. In this sense, the main issue with fake news lies not just in its inaccuracy but in its potential to fragment societies into separate epistemic communities, each with its own set of "facts". This fragmentation poses a significant threat to social cohesion and democratic governance, as it undermines the ability of societies to engage in informed and constructive dialogue (Abiri & Buchheim, 2022, p. 54).

Moreover, as Calil (2022, p. 176) outlines, fake news has a profound impact on public discourse and democratic processes. It can significantly influence public opinion and electoral outcomes by presenting misleading information about candidates or policies. This misinformation can sway voter perceptions and decisions, potentially altering the course of democratic processes. By distorting the truth, fake news undermines the integrity of elections and the democratic ideals of informed decision-making and accountability.

Addressing these symptoms is the central concern of this research, as societal coherence demands the legitimacy of its governing institutions, both political and scientific. The scientific legitimacy is arguably even more crucial, as such knowledge should guide best political practices. When scientific facts are disputed or misrepresented, it becomes exceedingly difficult to formulate policies based on sound evidence, further eroding public trust and effective governance.

3. FAKE NEWS: REGULATORY FRAMEWORK SCHEMES

3.1. Germany / Europe

Internet regulation in Europe is predominantly guided by a "notice-and-takedown" approach, prominently represented by Germany's Network Enforcement Act (*Netzwerkdurchsetzungsgesetz* – NetzDG)[5] and a 'notice-and-action' approach found in the EU's Digital Services Act (DSA).[6]

Germany's NetzDG, effective from October 1, 2017, epitomizes the world's principal Internet regulatory model. It aims to enhance digital intermediaries' efforts to address problematic online content by mandating a regulatory framework with severe penalties for non-compliance. NetzDG has sparked controversy and concern over its impact on freedom of speech and fundamental rights both within Germany and internationally (Stephenson & Rinceanu). NetzDG employs a "notice-and-takedown" approach that requires extensive public and private cooperation. Digital media platforms must delete or block illegal content within specified timeframes, ranging from 24 hours to seven days.[7] Illegal content is defined by various infractions in Germany's Criminal Code, including offences such as insult and public order disturbances.[8] Platforms are obligated to inform complainants

[5] Act to Improve Enforcement of the Law in Social Networks *(Netzwerkdurchsetzungsgesetz* - NetzDG*), Federal Law Gazette* I at 3352, enacted 1 October 2017.

[6] European Union**,** Regulation (EU) 2022/2065 on a Single Market for Digital Services (Digital Services Act - DSA) and amending Directive 2000/31/EC, [2022] OJ L277/1, entered into force on 16 November 2022.

[7] § 3(2)(2): "Social networks must delete or block access to manifestly unlawful content within 24 hours of receiving a complaint." § 3(2)(3): "For all other unlawful content, the deadline for deletion or blocking access is seven days after receiving the complaint."

[8] Criminal Code *(Strafgesetzbuch* - StGB*)*, last amended by Article 1 of the Law of 28 March 2023, *Federal Law Gazette* I p. 368. § 185 StGB - Insult (*Beleidigung*); § 186 StGB - Defamation (*Üble Nachrede*); § 187 StGB - Malicious Gossip (*Verleumdung*); § 130 StGB - Incitement to Hatred (*Volksverhetzung*); § 201 StGB - Violation of the Privacy of the Spoken Word (*Verletzung der Vertraulichkeit des Wortes*); § 201a StGB - Violation of Privacy through Taking Unauthorized Photographs (*Verletzung des höchstpersönlichen Lebensbereichs durch Bildaufnahmen*);§ 202a StGB - Data Espionage (*Ausspähen von Daten*); § 202b StGB - Phishing/Interception of Data (*Abfangen von Daten*); § 202c StGB - Preparation of Unauthorized Data Access (*Vorbereiten des Ausspähens und Abfangens von Daten*); § 241 StGB - Threat (*Bedrohung*), among others.

of their decisions and any rights of appeal[9] and report their content moderation activities publicly.[10] Additionally, platforms must report potentially criminal content, including IP addresses, to Germany's Federal Criminal Police Office (*Bundeskriminalamt*), notifying users no earlier than 4 (four) weeks after this transmission.[11] Non-compliance can result in fines of up to €50 (fifty) million for corporations and €5 (five) million for corporate officials.[12] The popularity of Germany's regulatory approach is evident, with over 25 countries and the EU adopting or proposing legislation influenced by NetzDG.[13]

The EU's Digital Services Act (DSA) based on a 'notice-and-action' model/approach is a significant testament to the influence of Germany's NetzDG. Enacted to shape Europe's digital future, the DSA aims to create a safe, predictable, and trustworthy online environment by countering harmful content such as hate speech, disinformation, and other objectionable content, while upholding fundamental rights.

Directly applicable to all 27 EU Member States, the DSA places primary responsibility on EU-based private digital intermediaries for handling illegal online content. Similar to NetzDG's "notice-and-takedown" model, the DSA introduces a "notice-and-action" mechanism, requiring digital platforms to provide an accessible procedure for users to report illegal content. The DSA defines "illegal content" broadly in Art. 3(h) as any information not compliant with Union law or the law of any Member State compliant with Union law.[14] This definition is broader than the German counterpart, which covers only violations of designated criminal provisions, as previously mentioned.

Complaints about illegal content can be submitted by individuals or entities, and platforms must respond in a timely, diligent, non-arbitrary, and objective manner, notifying complainants of decisions and legal remedies.[15] Notices from "trusted flaggers" receive priority and expedited processing; "trusted flagger" status is granted to entities with expertise in handling illegal content, such as Europol and the INHOPE Association.[16] Additionally, Art. 9 of the DSA requires platforms to comply with EU Member State orders to act against specific illegal content.

The DSA differs from NetzDG in several key ways. First, it does not prescribe specific timeframes for content removal, allowing platforms flexibility to make timely decisions and exempting them from liability if they act diligently.[17] Platforms must explain to users any restrictions imposed and their legal or contractual basis, with options for users to appeal through internal mechanisms, out-of-court settlements, or judicial redress.[18] Second, unlike

[9] § 3(2) of the NetzDG

[10] § 3(5) of the NetzDG

[11] § 3a NetzDG

[12] § 4 NetzDG

[13] *Search results: NETZGD*. Available at: https://justitia-int.org/?s=NetzDG (10. 10. 2024).

[14] Art. 3(h) DSA.

[15] Arts. 16, 17, 18 DSA.

[16] Art 22 DSA.

[17] Art 14 DSA.

[18] Art 20 DSA.

NetzDG's strict requirements, the DSA mandates that platforms notify authorities only when they suspect a criminal offence involving a threat to life or personal safety.[19] Third, the DSA does not require platforms to continuously monitor website traffic for illegal content.[20]

In conclusion, the intent behind the Digital Services Act (DSA) is indeed to create a cohesive regulatory framework across the European Union, addressing illegal content and establishing clear rules for digital platforms. This framework aims to replace or supersede fragmented national regulations, such as Germany's Netzwerkdurchsetzungsgesetz (NetzDG), thereby achieving consistency and uniformity within the EU's single market.[21]

It is crucial to observe, that the European Union's Digital Services Act (DSA) has attracted criticism for its potential to lead to over-censorship and for perceived gaps in its enforcement mechanisms. While the Act aims to tackle disinformation and illegal content, critics worry that granting the European Commission direct intervention powers during crises could prompt hasty or excessive content removal, affecting lawful expressions, such as satire and political critique. This concern is amplified by recent proposals that would allow the Commission to unilaterally declare a crisis and dictate platforms' responses, which could undermine the careful balance initially sought by lawmakers to protect free expression while combating disinformation.[22]

Furthermore, the DSA's transparency provisions, though beneficial for accountability, include exceptions that could allow platforms to withhold information deemed sensitive, potentially limiting the Act's effectiveness. Critics also highlight that while the DSA permits vetted researchers access to data for compliance assessments, trade secret protections might restrict meaningful analysis, weakening external oversight. Additionally, the DSA's "light-touch" approach to disinformation allows platforms significant discretion over "lawful but awful" content, leading to concerns about inconsistent moderation practices and the Act's overall ability to manage disinformation at scale effectively.[23]

3.2. United States of America

The United States employs a "market self-regulation" model, arguably symbolizing deep disagreement on the constitutional role of freedom of expression in democratic nations. The primary regulatory framework in Section 230 of the Communications

[19] Art. 21 DSA.

[20] Art. 7 DSA.

[21] European Parliament and Council Regulation (EU) 2022/2065 of 19 October 2022 on a Single Market for Digital Services (Digital Services Act) and amending Directive 2000/31/EC [2022] OJ L277/1, recitals 9, 10 and 14.

[22] Meyers, Z. 2022. *Will the Digital Services Act save Europe from disinformation?* Centre for European Reform. Available at: https://www.cer.eu/insights/will-digital-services-act-save-europe-disinformation (9. 10. 2024); Buri, I. & van Hoboken, I. 2021. *The DSA proposal: a critical overview.* Institute for Information Law, DSA Observatory. Available at: https://dsa-observatory.eu/wp-content/uploads/2021/11/Buri-Van-Hoboken-DSA-discussion-paper-Version-28_10_21.pdf (10. 10. 2024).

[23] Amnesty International. 2022. *What the Digital Services Act means for human rights and harmful Big Tech business models. Amnesty International EU Office.* Available at: https://www.amnesty.org/en/documents/pol30/5830/2022/en/ (10. 10. 2024); Buri, I. & van Hoboken, I. 2021. *The DSA proposal: a critical overview.* Institute for Information Law, DSA Observatory. Available at: https://dsa-observatory.eu/wp-content/uploads/2021/11/Buri-Van-Hoboken-DSA-discussion-paper-Version-28_10_21.pdf (10. 10. 2024).

Decency Act (CDA)[24], which protects digital platforms from civil liability for offensive speech acts.[25] Courts have interpreted this broadly to protect against claims based on third-party content, including negligence, deceptive trade practices, unfair competition, and more.[26] This broad safe harbour is considered essential for a functioning Internet. This provision has been broadly interpreted by courts to shield digital platforms from liability for content created by users. This includes protection from claims of negligence, deceptive trade practices, unfair competition, and more. The broad safe harbour provision is considered essential for maintaining a functional Internet, as it allows platforms to host user-generated content without fear of constant litigation.[27]

It is relevant to mention the current state of Litigation and Legislative Responses as numerous issues related to content moderation and free speech are currently being litigated before the US Supreme Court in cases such as Moody v. NetChoice, LLC.[28] Over 100 bills have been proposed in state legislatures to regulate social media platforms' content moderation policies. For instance, Florida's Senate Bill 7072[29] sought to regulate social media platforms by requiring transparency in censorship decisions and consistent application of standards. However, the Eleventh Circuit Court of Appeals declared it unconstitutional, raising critical questions about the nature of digital platforms' roles as "speech" or "editorial discretion" and whether they can be regulated as "common carriers"[30].

In summary, Section 230 provides significant protection for digital platforms, including from liability for hosting offensive speech, such as hate speech, aligning with First Amendment principles.[31] Although such content is protected by free speech laws, platforms address it by setting their content policies or updating their Terms of Use to balance user expression with community standards. The shift from a traditional "broadcasting" model to an interactive "participatory" model has positioned these platforms as new gatekeepers. This role raises tensions between their profit-driven business models and their responsibilities to uphold human rights, such as freedom of expression, privacy, and protection from harm

[24] United States, Communications Decency Act, 47 U.S.C. § 230 (1996).

[25] Section 230 The text of Section 230(c)(1) states: "No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider."

[26] Cases on the broad interpretation of Section 230: *Zeran* v. *America Online, Inc*., 129 F.3d 327 (4th Cir. 1997); *Fair Housing Council of San Fernando Valley* v. *Roommates.com,* LLC, 521 F.3d 1157 (9th Cir. 2008); *Jones* v. *Dirty World Entertainmen*t *Recordings* LLC, 755 F.3d 398 (6th Cir. 2014); *Doe* v. *MySpace, Inc*., 528 F.3d 413 (5th Cir. 2008); *Force* v. *Facebook, Inc*., 934 F.3d 53 (2d Cir. 2019).

[27] Importance of safe harbour for a functioning Internet see further: Vogus, C. 2021. *Answers to Five Key Questions from House Energy & Commerce Section 230 Hearing.* Center for Democracy and Technology. Available at: https://cdt.org/insights/answers-to-five-key-questions-from-house-energy-commerce-section-230-hearing/ (10. 10. 2024).

[28] *Moody v. NetChoice*, LLC, 603 U.S. (2024)

[29] Florida Senate Bill 7072, 2021; *Governor Ron DeSantis Signs Bill to Stop the Censorship of Floridians by Big Tech.* 2021. Available at: https://www.flgov.com/2021/05/24/governor-ron-desantis-signs-bill-to-stop-the-censorship-of-floridians-by-big-tech/ (10. 10. 2024).

[30] *NetChoice, LLC, et al.* v. *Attorney General, State of Florida, et al*., No. 21-12355 (11th Cir. 2022).

[31] United States Constitution, Amendment I.

Digital platforms thus navigate a complex landscape where they must reconcile their commitment to free speech with pressures to moderate content. Their policies and practices influence public discourse, raising questions about transparency, accountability, and the balance between enabling open dialogue and restricting harmful content. This delicate position often places platforms at the centre of debates about the limits of Section 230 and the role of private entities in regulating speech in digital spaces.

3.3. Canada

Canada has adopted a distinctive 'hybrid' model of online governance, reflected in its proposed Bill C-63,[32] which pivots from traditional "notice-and-takedown" systems to a more nuanced "systems-based" regulatory approach. This approach emphasizes collaboration among various stakeholders and aims to address the complexities of modern digital communications.[33]

Therefore, Canada's new regulatory proposal, embodied in Bill C-63, marks a shift from conventional content removal strategies to a systems-based risk assessment model. This model mandates a "duty to act responsibly" for digital platforms, focusing on transparency and systemic decision-making processes upstream of conventional content review mechanisms.[34]

Bill C-63 emphasizes proactive risk management, requiring digital platforms to implement measures to mitigate harmful content before it escalates.[35] The Canadian government engaged in extensive public consultations to inform this regulatory framework. Input from citizens and experts highlighted concerns about potential overreach and privatized censorship, emphasizing the need for precise definitions of harmful content, caution against proactive monitoring, and transparency in enforcement actions.[36]

A significant aspect of the Canadian model is its attempt to balance regulatory actions with the protection of free expression. Expert consultations underscored the importance of not incentivizing general monitoring, which could lead to over-censorship and infringe on free speech rights.[37] Canada's approach involves a diverse array of stakeholders, including public and private entities, fostering a more holistic regulatory environment. This multi-stakeholder approach integrates socio-technical-legal elements into the regulatory framework, ensuring a broader perspective on digital governance.[38]

Drawing insights from fields such as medical diagnostics and social medicine, Canada's model incorporates systemic causation and contextual regulatory measures, enhancing the effectiveness and adaptability of online governance strategies. This

[32] Online Harms Act, Bill C-63, 1st Sess, 44th Parl, 2024.

[33] Rinceanu, J. & Stephenson, R. 2024. *Differential Diagnosis in Online Regulation*. Eucrim. Available at: https://eucrim.eu/articles/differential-diagnosis-in-online-regulation/ (10. 10. 2024).

[34] *Online Harms Act, Bill C-63*, 1st Sess, 44th Parl, 2024, s. 3.

[35] s. 5.

[36] s. 7.

[37] s. 11.

[38] s. 15.

interdisciplinary approach aims to create a resilient and adaptable framework capable of addressing the evolving challenges in the digital landscape.[39]

In conclusion, the Canadian regulatory model exemplified by Bill C-63, the Online Harms Act, underscores a balanced and adaptive approach to online safety. Through extensive public and expert consultations, the government integrated diverse perspectives to address systemic factors and emphasize transparency in regulatory practices.[40] By moving beyond conventional content removal, this systems-based framework imposes a "duty to act responsibly" on digital platforms, promoting proactive risk management. The approach mitigates risks such as overreach and privatized censorship, ensuring accountability while preserving free expression. This robust framework not only strengthens the regulatory landscape but also adapts to the evolving challenges of online harms in a way that aligns with democratic values and public expectations.[41]

4. A HOLISTIC AND INTEGRATED APPROACH TO THE REGULATORY REGIME OF TECH GIANTS

4.1. The Need for a Global Regime

The regulation of fake news, hate speech, and other harmful online content varies substantially across regions, resulting in a fragmented and often ineffective global landscape, as previously described. These regulatory approaches differ so drastically that achieving a unified global framework appears nearly impossible, compounded by the unique limitations and challenges each system faces.

Firstly, they often treat issues like fake news and hate speech as isolated problems, leading to fragmented and ineffective regulations. For instance, the EU Digital Services Act tends to address these issues separately, failing to consider the interconnected nature of the digital media landscape.

Secondly, the varying regulations adopted by different countries can lead to unintended consequences, such as censorship or the spread of propaganda, due to a lack of contextual adaptation. The differing approaches between Europe and the USA highlight this issue, resulting in inconsistencies in enforcement and effectiveness. The lack of a unified global approach complicates the enforcement of consistent anti-misinformation measures, creating a fragmented regulatory environment (Abiri & Buchheim, 2022).

Thirdly, these regulatory approaches are primarily reactive, dealing with fake news after it has already been disseminated. Despite extensive fact-checking efforts during the USA 2020 elections and the COVID-19 pandemic, misinformation continued to significantly influence public perception (Humprecht, 2018).

[39] s. 20.

[40] Government of Canada. 2021. *Have your say: The Government's proposed approach to address harmful content online*. Available at: https://www.canada.ca/en/canadian-heritage/campaigns/harmful-online-content.html (10. 10. 2024).

[41] For further details on this framework and its impact on digital platforms, refer to: Salloum, J. *et al.* 2024. *Canada's new Online Harms Act (C-63): what you need to know.* Osler. Available at: https://www.osler.com/en/insights/updates/canada-s-new-online-harms-act-c-63-what-you-need-to-know/?pdf=1 (10. 10. 2024).

Fourthly, aggressive content moderation raises concerns about censorship and the suppression of free speech. Efforts to curb misinformation must be carefully balanced to avoid infringing on individual rights to freedom of expression. Overzealous content removal can stifle legitimate discourse and contribute to perceptions of bias and unfairness, further eroding public trust in digital platforms (Calil, 2022). Variations in human rights protections and constitutional structures pose significant challenges, particularly in the context of filtering and blocking online speech. Effective regulation requires a nuanced understanding of political and constitutional contexts (Humprecht, 2018).

Fifthly, the immense power of private digital platforms, which own and control much of the Internet's infrastructure, facilitates privatized government censorship. This, combined with economic incentives, threatens the quality and quantity of public discourse. The economic interests of these platforms often conflict with the public's need for reliable information, leading to a privatized form of censorship that undermines democratic processes (Calil, 2022).

Sixthly, addressing the root causes of misinformation requires regulatory frameworks that consider the underlying economic incentives and the role of algorithmic amplification by digital platforms (Abiri & Buchheim, 2022). These economic incentives drive the spread of fake news, making it profitable to create and disseminate sensational stories that attract clicks and generate ad revenue.

Finally, the analysis of existing regulatory frameworks reveals significant deficiencies in addressing the complexities of the digital media landscape. The pervasive issue of digital fake news poses a substantial threat to democracies, public health, and even the future of our planet. Despite efforts such as fact-checking and content removal during the USA 2020 elections and the COVID-19 pandemic, a significant portion of the population continues to believe in misinformation. This indicates that truth-based solutions like fact-checking do not fully address the problem. To effectively combat misinformation and protect democratic integrity, there is a critical need for a unified global regulatory regime. This regime should be based on comprehensive principles, norms, and rules that are adaptable to various socio-political contexts, ensuring consistent enforcement and effectiveness across different regions. It should integrate proactive strategies, technological solutions, and interdisciplinary approaches to create a balanced and effective regulatory framework. Only through such a holistic and integrated approach can we hope to address the root causes of misinformation and foster a healthier digital public sphere.

4.2. Regulatory Regime Basis and Structure

The regulation of fake news, hate speech, and other harmful online content can benefit from methodologies inspired by social medicine and comparative law. Drawing parallels between medical diagnosis (as it was previously suggested in Section 2.2) and legal regulation, we argue for a more holistic and integrated approach. This approach acknowledges the multifaceted nature of the digital media landscape and incorporates various stakeholders to create a comprehensive strategy (Flew, Martin & Suzor, 2019).

Historically, the perspective of Rudolf Virchow, a 19th-century physician, underscores the importance of considering social determinants in addressing health issues. Virchow advocated that political actions are necessary to address societal health problems, an approach that can be analogously applied to digital misinformation, where societal factors play a crucial role (Taylor & Rieger, 1985). Similarly, George Engel's biopsychosocial model from the 1960s integrates biological, psychological, and social factors in understanding health and illness. This model emphasizes the interconnected nature of these factors, relevant to the digital domain where technological, psychological, and social elements are deeply intertwined (Engel, 1977; Stephenson & Rinceanu, 2023, p. 75).

Comparative legal methodology, particularly functionalism, offers valuable insights into regulatory frameworks. This method aims to uncover broader socio-political connections underlying legal doctrines by gathering and interpreting information about various legal systems, evaluating similarities and differences between domestic legal regimes, and developing hypotheses to address shared regulatory challenges (Zweigert & Kötz, 1998). Such an approach ensures that regulatory measures are context-sensitive and adaptable to different socio-political landscapes.

Using this interdisciplinary methodological approach within a comparative legal framework, we propose a rational reconstruction of the regulatory regime for tech giants. When John Ruggie introduced the concept of international regimes into the international politics literature in 1975, he defined a regime as: "a set of mutual expectations, rules and regulations, plans, organizational energies and financial commitments, which have been accepted by a group of states" (Ruggie, 1975, p. 570).

Later, Krasner (1983, p. 2) elaborated on this definition, describing international regimes as: "sets of implicit or explicit principles, norms, rules and decision-making procedures around which actors' expectations converge in a given area of international relations" (Krasner, 1983, p. 2). Principles are beliefs about a fact, cause, or reaction. Norms are standards of behaviour defined in terms of rights and obligations. Rules are specific prescriptions and prospects for action. Decision-making processes establish actions that tend to prevail, practices for making and implementing collective choices.

"The principles of the regime generally define the objectives expected to be pursued by its members." (Krasner, 1983, p. 4). For example, the norms of the Paris Agreement (2015) do not require its members to immediately achieve all climate goals but incorporate prescriptions for members to practice transparency and accountability, aiming to lead them towards gradual and sustained climate action." For a regulatory regime addressing fake news and harmful content, principles would involve commitments to transparency, accountability, and the protection of human rights. These principles should harmonize the need for free expression with the imperative to prevent harm caused by misinformation.

Ronald Dworkin's theory (Dworkin, 1977) emphasizes the importance of principles over mere rules. The author argues that the community's political practices should express principles that go beyond simple rules. He distinguishes between three models of political association: the associative model, the rules model, and the principles model (Dworkin, 1977). The principles model is particularly relevant as it insists that members of a political community are genuinely linked by common principles, not just by rules created through political agreements. This model satisfies the conditions of a pluralistic society by ensuring that citizens respect the principles of their particular community, even if these differ from those of other communities.

Norms, in the context of Dworkin's framework, serve as a bridge between principles and rules. They are standards of behaviour that carry the weight of principles but provide more specific guidance similar to rules. In the regulatory regime for tech giants, norms would include standards for content moderation, such as the duty to remove or flag misinformation and the obligation to ensure that such actions do not unjustly infringe on freedom of expression. These norms must be adaptable to various socio-political contexts, allowing for effective implementation across different legal landscapes (Humprecht, 2018, p. 10). Dworkin's distinction between principles, norms, and rules highlights that norms carry weight and importance, providing the flexibility needed to adapt to specific contexts while being grounded in overarching principles of justice and equity. Norms in this regulatory regime ensure that the actions of tech giants align with the broader principles of the regime, providing a balanced approach to regulation. Arguably allowing the international community to enshrine the set principles on legal agreements serving as a basis for the rules to be developed in domestic regulation schemes.

Rules are more specific than norms, detailing the rights and obligations of regime members. For tech giants, rules would specify procedures for content removal, the steps required to verify the authenticity of content, and the obligations to provide users with mechanisms to appeal content moderation decisions. Rules should also encompass requirements for transparency reports and data sharing with regulatory bodies to enhance accountability (Abiri & Buchheim, 2022, p. 110).

Dworkin argues that the application of rules requires discretion and judgment, especially in "hard cases" where rules may conflict or be insufficient (Dworkin, 1977). In such instances, judges and regulators must resort to principles to guide their decisions, ensuring that the outcomes align with the broader principles of justice and equity.

As for the decision-making processes in this regulatory regime should promote the implementation of principles, norms, and rules through collaborative and interdisciplinary approaches. This includes engaging stakeholders from governments, private digital platforms, civil society, and international organizations. Educational initiatives are essential, promoting digital literacy to help users critically assess information (Caled & Silva, 2022, p. 135). Leveraging advanced technologies such as artificial intelligence and machine learning can enhance the detection and mitigation of fake news, making content moderation more effective (Stephenson & Rinceanu, 2023, p. 79).

Effective regulation must also consider the unique political, cultural, and legal contexts of different regions. Disinformation strategies vary significantly across countries, influenced by national news agendas and political cultures, necessitating tailored regulatory approaches (Humprecht, 2018, p. 85).

As a result of the suggested interdisciplinarity approach it is feasible to incorporate proactive strategies that are necessary to prevent the spread of misinformation. This includes real-time monitoring and early intervention mechanisms to address fake news before it gains traction. Addressing the root causes of misinformation, such as economic incentives and algorithmic amplification, is vital for a sustainable solution (Abiri & Buchheim, 2022, p. 95).

In summary, the ideas previously developed constitute the theoretical basis for a Global Regulatory Regime for Tech Giants. This framework aligns with the format set by Sabino Cassese, integrating core legal structures with soft law and interdisciplinary aspects. Utilizing tools from Comparative Law, this approach aims to create a balanced and effective regime that not only holds tech giants accountable but also fosters a trustworthy information ecosystem (Cassese, 2005, p. 47).

Implementing this integrated approach requires collaboration among various stakeholders, including governments, private digital platforms, civil society, and international organizations. Global regulatory systems, as articulated by Cassese, thrive on mutual connections and joint decision-making processes. This involves the active participation of states, sub-state entities, and international bodies to create a cohesive regulatory environment (Cassese, 2005, p. 45). Such an environment ensures consistent enforcement and effectiveness across different regions, addressing the root causes of misinformation and promoting a healthier digital public sphere.

By leveraging Cassese's insights (Cassese, 2005) into global regulation, the proposed framework offers a practical pathway to operationalize the interdisciplinary and harm-based approaches discussed. This comprehensive strategy not only addresses the immediate impacts of fake news but also seeks to understand and mitigate the underlying conditions that allow such misinformation to proliferate. Through fostering transparency, accountability, and inclusive dialogue, this regime aims to restore the common factual ground necessary for democratic legitimacy and social cooperation, ultimately enhancing societal trust and resilience in the digital age.

5. CONCLUSION

The pervasive issue of fake news poses a significant threat to the integrity of information disseminated online, particularly by tech giants. This paper has examined the regulatory mechanisms that hold these corporations accountable for the spread of misinformation and evaluated their effectiveness as deterrents. Our analysis reveals that current regulatory frameworks, which often focus on symptom treatment such as content removal and fact-checking, are insufficient for addressing the root causes of misinformation.

The harm-based approach proposed in this study advocates for a holistic and integrative regulatory regime that goes beyond merely reacting to false content. By drawing on methodologies from social medicine and comparative law, we emphasize the importance of understanding the socioeconomic incentives, psychological drivers, and technological dynamics that fuel the production and dissemination of fake news.

Key to this approach is the development of robust legal frameworks that mandate transparency and accountability from tech giants. This includes distinguishing the responsibilities of content creators and sharers, particularly in relation to public agents, to ensure that accountability is appropriately distributed. The integration of soft law and guidelines can further promote ethical behaviour and best practices among digital platforms, fostering a culture of responsibility.

Drawing on Ronald Dworkin's distinction between principles, norms, and rules, this paper advocates for a regulatory regime that incorporates these elements to create a coherent and adaptable framework. Principles provide the foundational values, norms serve as guidelines for behaviour, and rules specify the rights and obligations of regime members. This structure ensures that the regulatory measures are grounded in justice and equity, adaptable to various socio-political contexts, and capable of addressing the underlying issues of misinformation.

The proposed harm-based regulatory regime aims to proactively combat the spread of fake news by fostering a more resilient and trustworthy digital information ecosystem. By focusing on the root causes and integrating legal, technological, and educational measures, this approach provides a balanced and effective solution to the pervasive issue of fake news.

In conclusion, the need for a holistic and integrative regulatory regime based on a harm-based approach is paramount in addressing the multifaceted nature of fake news. Such a regime not only enhances the resilience of digital information ecosystems but also ensures the restoration of societal trust and the protection of democratic integrity. Through comprehensive and proactive measures, this approach can significantly mitigate the impact of misinformation and foster a healthier digital public sphere.

By drawing on the concepts of global regulatory systems, differentiating responsibilities, promoting digital literacy, and leveraging technology, the proposed regime can effectively address the challenges of digital misinformation. Moreover, by focusing on reparation and the restoration of societal trust, the regime can foster a more trustworthy and reliable information ecosystem.

This comprehensive strategy ensures that regulatory measures are context-sensitive and adaptable to different socio-political landscapes, promoting ethical behaviour online and fostering a healthier digital public sphere. Through proactive and collaborative efforts, this approach can significantly mitigate the impact of misinformation and strengthen the resilience of digital information ecosystems.

ACKNOWLEDGEMENTS

I extend my deepest gratitude to Prof. Dr Fernando Menezes de Almeida and Dr Rodrigo Garcia Cadore, LL.M., for their invaluable guidance throughout the Comparative Constitutional Practices course at the University of São Paulo. Their expertise has been instrumental in shaping this paper. I also thank my colleagues for their insights, particularly during the session on Practices of Constitutional Emotionalization and Rationalization (24 November 2023), which focused on Fake News and Trust in Supreme and Constitutional Courts. The discussions with guest discussant Ana Luíza Calil, and the works of Edda Humprecht and Gulad Alibri/Johannes Buchheim, have greatly enriched my understanding.

LIST OF REFERENCES

Books and Articles

Abiri, G. & Buchheim, J. 2022. Beyond True and False: Fake News and the Digital Epistemic Divide. *Michigan Technology Law Review*, 29, pp. 59-109. https://doi.org/10.36645/mtlr.29.1.beyond

Amnesty International. 2022. *What the Digital Services Act means for human rights and harmful Big Tech business models. Amnesty International EU Office*. Available at: https://www.amnesty.org/en/documents/pol30/5830/2022/en/ (10. 10. 2024).

Buri, I. & van Hoboken, I. 2021. *The DSA proposal: a critical overview.* Institute for Information Law, DSA Observatory. Available at: https://dsa-observatory.eu/wp-content/uploads/2021/11/Buri-Van-Hoboken-DSA-discussion-paper-Version-28_10_21.pdf (10. 10. 2024).

Caled, D. & Silva, M. J. 2022. Digital media and misinformation: An outlook on multidisciplinary strategies against manipulation. *Journal of Computational Social Science*, 5, pp. 123–159. https://doi.org/10.1007/s42001-021-00118-8

Calil, A. L. 2022. Public Agents in Social Media Regulation: The Brazilian Case in a Comparative Perspective, *Journal of Law, Market & Innovation*, 1(2), pp. 162-182.

Cassese, S. 2005. Administrative Law Without the State? The Challenge of Global Regulation. *New York University Journal of International Law and Politics*, 37(4), pp. 663-684.

Dworkin, R. 1977. *Taking Rights Seriously*. London: Duckworth.

Engel, G. L. 1977. The Need for a New Medical Model: A Challenge for Biomedicine. *Science*, 196(4286), pp. 129-136. https://doi.org/10.1126/science.847460

Flew, T., Martin, F. & Suzor, N. 2019. Internet Regulation as Media Policy: Rethinking the Question of Digital Communication Platform Governance. *Journal of Digital Media & Policy*, 10(1), pp. 33-50. https://doi.org/10.1386/jdmp.10.1.33_1

Government of Canada. 2021. *Have your say: The Government's proposed approach to address harmful content online*. Available at: https://www.canada.ca/en/canadian-heritage/campaigns/harmful-online-content.html (10. 10. 2024).

Governor Ron DeSantis Signs Bill to Stop the Censorship of Floridians by Big Tech. 2021. Available at: https://www.flgov.com/2021/05/24/governor-ron-desantis-signs-bill-to-stop-the-censorship-of-floridians-by-big-tech/ (10. 10. 2024)

Humprecht, E. 2018. Where Fake News Flourishes: A Comparison across Four Western Democracies. *Information, Communication & Society*, 22(13), pp. 1973-1988. https://doi.org/10.1080/1369118X.2018.1474241

Krasner, S. D. 1983. International Regimes. In: Krasner, S. D. (ed.), *International Regimes*. Ithaca, NY: Cornell University Press, pp. 1-21.

Meyers, Z. 2022. *Will the Digital Services Act save Europe from disinformation?* Centre for European Reform. Available at: https://www.cer.eu/insights/will-digital-services-act-save-europe-disinformation (9. 10. 2024).

Salloum, J. *et al.* 2024. *Canada's new Online Harms Act (C-63): what you need to know.* Osler. Available at: https://www.osler.com/en/insights/updates/canada-s-new-online-harms-act-c-63-what-you-need-to-know/?pdf=1 (10. 10. 2024).

Max Planck Institute for the Study of Crime, Security and Law. 2024. *Rethinking Digital Media Regulation*. Available at: https://csl.mpg.de/en/projects/rethinking-digital-media-regulation?c=178896 (27. 6. 2024).

Rinceanu, J. & Stephenson, R. 2024. *Differential Diagnosis in Online Regulation*. Eucrim. Available at: https://eucrim.eu/articles/differential-diagnosis-in-online-regulation/ (10. 10. 2024).

Ruggie, J. G. 1975. International Responses to Technology: Concepts and Trends. *International Organization*, 29(3), pp. 557-583. https://doi.org/10.1017/S0020818300031696

Stephenson, R. & Rinceanu, J. 2023. Digital Iatrogenesis: Towards an Integrative Model of Internet Regulation. *Eucrim*, 1, pp. 73-80. Available at: https://eucrim.eu/articles/digital-iatrogenesis/ (10. 10. 2024).

Taylor, R. & Rieger, A. 1985. Medicine as a social science: Rudolf Virchow on the typhus epidemic in Upper Silesia. *International Journal of Health Services*, 15(4), pp. 547-559. https://doi.org/10.2190/XX9V-ACD4-KUXD-C0E5

Vogus, C. 2021. *Answers to Five Key Questions from House Energy & Commerce Section 230 Hearing.* Center for Democracy and Technology. Available at: https://cdt.org/insights/answers-to-five-key-questions-from-house-energy-commerce-section-230-hearing/ (10. 10. 2024).

Zweigert, K. & Kötz, H. 1998. *Introduction to Comparative Law.* Oxford: Clarendon Press; New York: Oxford University Press.

Legal Sources and Case-Law

Act to Improve Enforcement of the Law in Social Networks *(Netzwerkdurchsetzungsgesetz* - NetzDG*), Federal Law Gazette* I at 3352, enacted 1 October 2017

Criminal Code *(Strafgesetzbuch* - StGB*)*, last amended by Article 1 of the Law of 28 March 2023, *Federal Law Gazette* I p. 368.

Doe v. *MySpace, Inc*., 528 F.3d 413 (5th Cir. 2008)

European Parliament and Council Regulation (EU) 2022/2065 of 19 October 2022 on a Single Market for Digital Services (Digital Services Act) and amending Directive 2000/31/EC [2022] OJ L277/1.

Fair Housing Council of San Fernando Valley v. *Roommates.com*, LLC, 521 F.3d 1157 (9th Cir. 2008).

Florida Senate Bill 7072, 2021.

Force v. *Facebook, Inc*., 934 F.3d 53 (2d Cir. 2019).

Jones v. *Dirty World Entertainmen*t *Recordings* LLC, 755 F.3d 398 (6th Cir. 2014).

Moody v. NetChoice, LLC, 603 U.S. (2024)

NetChoice, LLC, *et al.* v. *Attorney General, State of Florida, et al.*, No. 21-12355 (11th Cir. 2022).

Online Harms Act, Bill C-63, 1st Sess, 44th Parl, 2024.

United States, Communications Decency Act, 47 U.S.C. § 230 (1996).United States Constitution.

United Nations (2015) Paris Agreement.

Zeran v. *America Online, Inc*., 129 F.3d 327 (4th Cir. 1997)