

Freedom of Expression in the Digital Ecosystem

Murilo César Ramos
University of Brasília
murilo@ccom.unb.br

Nélia Del Bianco
University of Brasília
neliadelbianco@gmail.com

Marcos Urupá
University of Brasília
marcosurupa@gmail.com

Bios

Murilo César Ramos is an Emeritus Professor at the School of Communication at the University of Brasília, founder, and Senior Researcher at the Center for Policies, Law, Economy, and Communication Technologies (CCOM), and the Communication Policies Laboratory (LaPCom) at the same university. He holds a Ph.D. in Communication from the University of Missouri-Columbia, USA (1982). His principal work is considered to be the book *As Margens da Estrada do Futuro: Comunicações, Políticas e Tecnologias* (<https://eptic.com.br/wp-content/uploads/2014/12/MuriloRamos.pdf>), published openly on the web in 1999, alongside numerous other publications, books, and articles over a four-decade career.

Nelia Del Bianco is a professor in the Postgraduate Program at the University of Brasília and a Researcher at the Center for Policies, Law, Economy, and Communication Technologies (CCOM) at the same university. She holds a Ph.D. in Communication from the University of São Paulo with a postdoctoral internship at the University of Seville. She has authored dozens of articles and book chapters on communication policies, public broadcasting, technological changes in the media, the audiovisual communication ecosystem, and audiovisual communication infrastructure.

Marcus Urupá holds a Ph.D. and Master's degree in Communication and Society from the University of Brasília, with a research focus on the internet and public policies. He has a lato-sensu postgraduate degree from the Foundation School of Sociology and Politics - FESPSP in Management and Public Policies and holds degrees in journalism and law. He served as Director of the Paraense Broadcasting Foundation - Funtelpa, acting as Marketing Director, Director of TV Cultura do Pará, and the Cultura Portal. He has authored articles and research on public communication policies, digital communication, digital journalism, internet and telecommunications regulation, and digital rights.

ABSTRACT

Purpose: This paper addresses a critical issue that has characterized the digital ecosystem in recent years: the use of social media as a means of expression. By critically examining the business models practiced by digital platforms, particularly the extensive exploitation of personal data for behavioral advertising, including for political purposes, the research aimed to understand how these practices have significant impacts on democracy and the exercise of fundamental rights worldwide. This was done by first establishing a theoretical framework centered on the classical thought of John Milton, Thomas Paine, and John Stuart Mill, and Jürgen Habermas' concept of the public sphere. Secondly, on an empirical level, by examining three normative instruments designed to address this complex issue and the risks it imposes on the constitutional guarantee of freedom of expression.

Methodology/Approach/Design: The research focused on analyzing the main legislative initiative currently being discussed in the Brazilian Congress, Bill 2630/2020. This analysis included examining supplementary material from public hearings and contributions from societal entities. Additionally, the mentioned bill was compared to the Digital Services Act (DSA) and the Online Safety Act.

Findings: The results indicate a conceptual convergence of the proposed Brazilian legislation with the European DSA. This raises more questions than answers, not only because the European norm is just now being enforced but also because none of the three normative texts address the main question the research attempted to problematize: the business models put forth by digital platforms most likely to affect freedom of expression as a fundamental right.

Practical implications: The study aims to provide legislators and regulators with the greatest possible certainty that their actions will be supported by a conceptual and institutional framework that reinforces, despite the pressures technological changes will inevitably exert, the constitutional guarantees of freedom of expression.

Originality value: While the theme of freedom of expression, a constitutional cornerstone, may not allow for much conceptual originality, its encounter with the digital ecosystem requires creativity not to be swayed by the allure of the new at the expense of what has proven philosophically solid over centuries, nor by the normalization of commercial practices that threaten not only institutions but also individual intimacy and individuality.

Keywords: Freedom of Expression, Digital Ecosystem, Legislation, Regulation, Bill No. 2630/2020, Digital Services Act, Online Safety Act.

ACKNOWLEDGMENT

This work was supported by the National Telecommunications Agency (Anatel) [Grant Number TED Anatel No. 1/2022]

INTRODUCTION

When it emerged as a technological artifact in the second half of the 20th century and unfolded into an important communication instrument—the World Wide Web, created by Tim Berners-Lee—the internet brought renewed hope for communicational empowerment of the human person through unrestricted access to this new form of expression. This seemed like the radical realization of hope embedded, after years of debate within UNESCO, expressed in the seminal study *Many Voices, One World* published in 1980. Or like the hopes previously placed on broadcasting, radio, and television, but frustrated, at least partially, over the first 60 years of the 20th century.

Based on this premise, the objective of this study was to explore the hypothesis that political-economic forces might have disfigured the World Wide Web to such an extent that, particularly after 2016 and the U.S. presidential election, national states have been required to make political, normative, and regulatory interventions on freedom of expression, generating an environment of growing uncertainties, from theoretical to institutional. These uncertainties stem from the disruptive power and speed characterizing the technical development of the internet but especially from the exponential accumulation of political power in a very small number of platforms.

The study began with a brief discussion on how the idea of freedom of expression evolved since the Enlightenment, using central aspects of the thoughts of John Milton, John Stuart Mill, Thomas Paine, and contemporarily, Jürgen Habermas, and his founding concept of the public sphere. The goal was to shed light on the democratic, institutional, political-normative, and regulatory challenges surrounding the internet and the Web today.

Next, the study aimed to understand how the original proposal of the Web, a virtuous circle leading to the fullest exercise of individual and collective freedom of expression, transformed into a socio-cultural and political-economic environment characterized, in the second decade of the century, by what Oxford Dictionaries chose as the word of the year in 2017: post-truth. This term initially had a strong metaphorical force but was later overshadowed by another term with poorer heuristic quality, but which dominated the global debate on information and disinformation since then, fake news, as if the

theoretical-conceptual and institutional impasses surrounding freedom of expression were restricted mainly to journalism, and not inherent to a fundamental human right.

This analysis will be conducted through the examination of two legislations already in force in the European Union and the United Kingdom, respectively the Digital Services Act and the Online Safety Act, and Bill No. 2630/2020, currently under discussion in the Brazilian Congress, whose latest version is largely inspired by the DSA. The analysis will focus on how freedom of expression is treated and whether this fundamental freedom is preserved or not from the risks of being placed under even greater threats in the name of protecting it.

In the end, this article will refer to the Solid Project, explored in greater depth in the study concluding the research project on New Regulatory Challenges of the Digital Ecosystem, specifically in the report on the Scenarios of Regulatory Institutional Conformation of the Digital Ecosystem. This reference will highlight the extent to which the business model based on the massive commercial exploitation of personal data by companies such as Meta, Alphabet, Twitter, Microsoft, and Amazon has been responsible for the theoretical, conceptual, political-economic, political-normative, and regulatory siege to which the fundamental right of freedom of expression is subjected today, to the point of threatening its very existence.

ON FREEDOM OF EXPRESSION

John Milton (1608-1674) is an essential reference for anyone seeking to understand the place and centrality of freedom of expression for human agency and how it unfolded in history. His seminal work, *Areopagitica – A Speech for the Liberty of Unlicensed Printing to the Parliament of England*, is a powerful, precursor manifesto for the freedom to publish without prior censorship.

However, as noted by Seelaender (1991, p.197):

“[...] one would never understand the positions taken by Milton in *Areopagitica* without examining the connections between them and the writer’s Protestant religion. The tendency to value the individual, implicit in the idea of universal priesthood, certainly underpinned phenomena such as the emergence in Miltonian thought of an inclination to defend citizens’ liberties. By demystifying the figure of the cleric, the equalizing inspiration of the priesthood principle fostered, through the deep symbiosis between spiritual and temporal powers in Stuart England, a growing aversion to tyrants who disrespected freedom.”

Seelaender’s interesting observation allows this study to do what is not usually done: to consider Milton’s magnum opus, *Paradise Lost*, as an equally essential contribution to the defense of freedom of expression as human agency. In this poem, considered one of the most important in the Western canon since Classical Antiquity, Milton gives a new meaning to the story of Adam and Eve—not of disobedience and punishment, the original sin imagined by Saint Augustine, but of affirmation of freedom and choice, even in the presence of divine omnipotence.

In another exemplary work, the historian and literary critic Stephen Greenblatt (2018), in *The Rise and Fall of Adam and Eve*, says of Milton, at the end of two chapters dedicated to *Paradise Lost*:

“The following passage, at the end of the poem, is one of the most beautiful Milton wrote. The verses continue to express faith in divine Providence, but even more so in freedom, the freedom Milton believed God conferred on the first couple, the freedom that still belonged to all men. In its ending, *Paradise Lost* releases Adam and Eve from the story that gave them life and sees them walk together toward an uncertain future: ‘Some natural tears they dropped, but wiped them soon; the world was all before them, where to choose their place of rest, and providence their guide: they hand in hand with wandering steps and slow, Through Eden took their solitary way.’”

And Greenblatt concludes the chapter: “More than a thousand years after Augustine, Adam and Eve at last became kings.”

In other words, returning to the vein that fuels this section of a study centered on freedom of expression, John Milton did more than write *Areopagitica*, a manifesto born from the English state's attempt to censor the book he wrote defending divorce as a right. In *Paradise Lost*, he made a radical defense of individual freedom of conscience, will, and choice, without which no one would ever have anything to express.

While Milton was the first voice in the emergence of political liberalism to produce rationally grounded arguments about the primacy of free will, even before divine Providence, and the right to use this free will to express all kinds of arguments before any concrete authority, it was not yet a time when this issue referred to the idea and institution that would come to be known as the press, or journalism. This institution, which had evolved from commercial correspondence systems essential to the emergence of capitalist modernity during its mercantile phase, enabled by the cycle of navigations, assumed its journalistic form, central to the American and French Revolutions, respectively from 1765-1783 and 1789-1799.

Thomas Paine (1737-1809) and John Stuart Mill (1806-1873), both Englishmen like Milton, are the most recognized names among the principal architects of the idea of freedom of the press as essential to the fullest enjoyment of freedom of expression.

Paine was mainly a revolutionary, directly involved as a protagonist and interpreter in the American and French Revolutions, for which his books *Common Sense*, in the former, and *The Rights of Men*, in the latter, were of fundamental importance. In *Common Sense*, despite being English, he argued for the independence of the English colonies in America, earning recognition as one of the Founding Fathers of the United States, though not on the same level as George Washington, Thomas Jefferson, Alexander Hamilton, James Madison, Benjamin Franklin, John Adams, Samuel Adams, and John Jay. Historical records show that *Common Sense* was the most circulated pamphlet in the colonies during the revolutionary war, with an estimated print run nearing 500,000 copies (Zinn, 2005, p. 69-70).

In France, during the revolution, Paine published *The Rights of Men*, a guide to the principal Enlightenment ideas, always in the same pamphlet style as his work published in the United States, using popular language that would later mark the still nascent journalism. Paradoxically, Paine's only text on the press and journalism was an article for the *American Citizens* newspaper in 1806, where he clearly distinguished between the act of printing as a mechanical process, later industrial, and the freedom to print as an act of citizenship.

Considered one of the greatest, if not the greatest, English-language philosophers of the 19th century, John Stuart Mill made a fundamental contribution to the debate on freedom of expression and the press in the form of a dense essay entitled *On Liberty*. In this work, he addresses the nature and limits of the power to be legitimately exercised over the individual in society, without curtailing their freedom.

Regarding the theme of this study, in chapter II on freedom of thought and discussion, Mill (2011) expresses his conviction that society benefits from the free exchange of ideas and opinions, and that the press plays a crucial role in facilitating this exchange. Key points of his thought on press freedom, expressed throughout the essay, include:

a) **Promotion of truth:** For Mill, press freedom is essential to the discovery and dissemination of truth. He argues that even if an idea is initially rejected by the majority, it should still be allowed to be expressed and debated, as truth often emerges from the clash of conflicting opinions. b) **Protection against tyranny:** He saw press freedom as a bulwark against tyranny and oppression, hence the importance of allowing diverse opinions to be heard, helping to prevent power concentration in the hands of a few and protect against abuse of authority. c) **Stimulation of progress:** Mill believed a free press would stimulate intellectual and social progress, encouraging innovation, challenging orthodoxies, and promoting debate. He also believed society benefits from individuals' freedom to express their opinions and challenge prevailing beliefs and institutions. d) **Importance of public scrutiny:** He emphasized the role of the press in holding governments and other powerful entities accountable, arguing that public scrutiny and criticism help ensure transparency and prevent power abuses. e) **Limits to government interference:** While defending press freedom, Mill recognized limits, especially when it comes to preventing harm to others.

Mill (2011, p. 56) further asserts:

"The liberty of opinion and the liberty of expressing opinions are necessary for the mental well-being of humanity (on which all its other well-being depends), based on four distinct grounds, which we now briefly recap. Firstly, even if an opinion is silenced, that opinion may, for all we know, be true. To deny this is to assume our own infallibility. Secondly, though the silenced opinion may be

erroneous, it may contain a portion of truth, which is often the case; and since the general or prevailing opinion on any subject rarely or never constitutes the whole truth, it is only through the collision of adverse opinions that the remainder of the truth has any chance of being supplied. Thirdly, even if the prevailing opinion is not only true but the whole truth; unless it is vigorously and earnestly contested, and until this happens, it will be held as a prejudice by most of those who receive it, with little understanding or feeling of its rational grounds. And not only this, but, fourthly, the meaning of the doctrine itself will be in danger of being lost, or enfeebled, and deprived of its vital effect on the character and conduct: the dogma becoming a mere formal profession, ineffectual for good, but cumbering the ground and preventing the growth of any real and heartfelt conviction from reason or personal experience."

Habermas and the Structural Transformation of the Public Sphere

Having established this very brief canon illustrating the trajectory of the idea of freedom of expression and the press between the 17th and 19th centuries, now follows a critical look inspired by a seminal work of youth by one of the most important philosophers and sociologists of the 20th century, Jürgen Habermas: *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society*.

In Habermas's deep archaeology to reach the category he sought to critique—the public sphere in bourgeois society, without giving it an operational definition as he saw it as a historical process in dialectical transformation—one finds between the lines the genesis of the mass press and the structural change it would provoke in a debate sphere, whose subject would be the very "public bearing public opinion" (Habermas, 1984, p. 14).

As this political public sphere expanded, ideas circulating among cafes, salons, and dining circles began to circulate also and preferably in a literary press transformed into an opinion press, composed of various periodicals and pamphlets, each reflecting the desires of a particular faction.

The effects of this revolution on the press could not have been more radical. On one side, technical advances allowed old manual presses to be replaced by steam-powered mechanical presses, capable of printing newspapers in larger quantities at increasing speeds. On the other, the rapid growth of markets accelerated the corresponding urbanization process, with the central cities of capitalism growing in population, potential consumers of goods offered by manufacturers and commerce: medicines, clothes, agricultural implements, household utensils.

From this social, political, and economic dialectic, a new type of press emerged from the 1830s, driven by a logic very different from the political press of the revolutionary bourgeoisie, whose initial contours and definitive forms inspired Milton, Paine, and Mill's writings on freedom of expression and the press.

The most visible mark of this great transformation in the press, giving it the contours of a nascent cultural industry, was a newspaper called *The New York Sun*, founded by Benjamin Day, first circulated on the streets of the great American city on September 3, 1833. The hallmark of this newspaper was its price: 1 penny, or 1 cent, compared to the 10 or 15 cents that newspapers of opinion press usually cost readers or subscribers.

From hundreds of initial copies, the *Sun* grew to 5,000 daily copies in 1834, and 15,000 copies in 1836, followed by dozens of other newspapers in major U.S. cities, with similar developments in central European countries. In the U.S., by 1896, just over 50 years from the penny press's inception, the *New York World*, created by magnate William Randolph Hearst (inspiring Orson Welles' character Citizen Kane), and the *New York Journal*, founded by Hungarian immigrant Joseph Pulitzer, engaged in a symbolic circulation war, the former reaching 600,000 daily copies against Pulitzer's 430,000.

In the 19th century, the communications universe, still driven by virtually ancient paper and printing technologies, began moving, pushed by new technical developments: the first public telegraphic message, the gramophone, photogravure or photography, the first submarine telegraph cable, the first wired telephone message, wireless telephone messages. However, from 1895, with cinema's emergence, the moving image, and from 1906, when Fessenden broadcast the human voice over radio waves, human speech underwent perhaps its most radical inflection since writing's invention.

Cinema and radio marked the first 30 years of the 20th century, progressively taking the form of the most pervasive, at least until 1989, industry of our era—the cultural industry of increasingly influential symbolic goods, responsible for new sociability, causing profound transformations in contemporary notions of time and space, altering human relationships between individuals, within families, changing classic forms of politics, influencing actions in World War I battlefields, ultimately making communications the most incisive cultural, political, and economic phenomenon until then. These enormous transformations found their synthesis from 1923 with the invention of a new device, neither radio nor cinema: television, which decisively marked 20th-century Western culture. It brought new and complex challenges for freedom of expression and the press, given its supporting technology: the radio frequency spectrum, a common, finite good impossible to exploit without some state intervention.

Since the 1920s and 1930s, every technological turn in communications, especially social communication, renewed civilizational hopes for more relevant information, greater public sphere participation, and more democracy in politics and economics. However, the Critical Theory formulated from those decades by Habermas's professors and colleagues at the Frankfurt School, including Theodor Adorno, Max Horkheimer, Erich Fromm, Walter Benjamin, Herbert Marcuse, Felix Weil, and others, had already critically revisited these hopes, consolidated in Adorno and Horkheimer's *Dialectic of Enlightenment*, published in 1944. At the core of this foundational work, which first employed the term "Cultural Industry," lies a philosophical critique of Enlightenment categories of reason and nature, accompanied by a profound exposition and analysis of modern thought's history, transitioning from mythology to philosophy and science. It also includes an alert about how technical progress can be accompanied by ideological regression, reflecting on how culture can underpin fascist barbarism.

This context mirrors our own era's spirit. An era in which, as pointed out in the Introduction, a new technological artifact, the internet, emerged in the second half of the 20th century, first in its pure form, then transformed into the global-reaching World Wide Web, generating all kinds of civilizational hope, not much different from those imagined by Milton, Paine, and Mill. However, little more than 30 years later, it is subjected to critical scrutiny not much different from that studied by the Frankfurt School decades ago. Among these is the idea that a significant portion of what travels on the Web today carries signs of a return to fascist-like barbarism.

The following section analyzes European, British legislation, and a Brazilian bill—the Digital Services Act, the Online Safety Act, and Bill No. 2630/2022. These normative initiatives are largely a response to the ethical crisis triggered by the Cambridge Analytica scandal, revealing the ethical and normative challenges posed by digital platforms exploiting massive personal data collection. This crisis, rooted in the business models underpinning most Web-operating companies—massive personal data exploitation—extends beyond mere commerce to politics and electoral processes, blending commercial and political marketing.

In summary, the 2016 U.S. electoral process, which resulted in Donald Trump's victory, saw Cambridge Analytica, under the guise of academic research using the fantasy name Cambridge, illegally appropriating about 87,000 Facebook user profiles through a third-party app called "This is Your Digital Life," along with another 270,000 legally consenting users without knowing their friends' data would be collected without authorization. This scenario, whether minimally or substantially, was instrumental in Trump's election, amid accusations of Russian cyber influence, email leaks from Hillary Clinton's campaign associates, and disinformation factories in Macedonia.

By the end of 2016, the Oxford Dictionaries chose Post-Truth as their word of the year, with the world plunging into a maelstrom of disinformation driven by all kinds of directed messages, a maelstrom still swirling today, to which this study seeks to contribute within its scope and capacity.

At the core of the study, as detailed below, lies a thorough analysis of two existing legislations—the European Digital Services Act; the United Kingdom's Online Safety Act; and the Brazilian Bill No. 2630/2020, currently under discussion in the National Congress, dealing with freedom, responsibility, and transparency on the internet. The aim is to evaluate how these legislations protect the fundamental right to freedom of expression and the press, as discussed in section 2 of this study.

POLITICAL-NORMATIVE AND REGULATORY INTERVENTIONS: DSA, ONLINE SAFETY ACT, AND BILL NO. 2630/2020

Digital Services Act (DSA)

Over the past two decades, digital platforms have become integral to daily life—it's hard to imagine doing anything online without Amazon, Google, or Facebook. For too long, these tech giants have benefited from a lack of rules, defining their content moderation policies and issuing "transparency reports" on their efforts to combat harms like disinformation.

Without regulation, these digital platforms have gained a dominant position over competitors, exerting undue influence on democracy, fundamental rights, societies, and the economy (European Parliament, 2023).

The European Union decided to address this issue with a set of regulations forcing major internet platforms to combat the spread of illegal content and other social risks under the threat of hefty fines. One such instrument is the Digital Services Act (DSA), published in the EU's official journal in October 2022 alongside the Digital Markets Act (DMA), focusing on restoring competitiveness in the digital sector.

The DSA's application was divided into two phases. In the first phase on August 25, 2023, the largest platforms and search engines operating in Europe were required to comply with legal governance provisions, transparency about their algorithmic systems, and accountability for social risks arising from their use. The second phase on February 17, 2024, extended the law's full applicability to all intermediary services covering the EU.

The DSA culminates several years of the EU's struggle against the spread of illegal online content, toxic content, and disinformation. In 2018, the European Commission issued a Code of Practice for online platforms, trade associations, and major advertising sector players, committing them to reduce disinformation and improve their policies. This self-regulation proposal garnered 16 signatories and 21 commitments on issues like thorough advertising channel analysis, transparency of political and issue-based advertising, service integrity, consumer empowerment, and fact-checking.

The European Commission evaluated this document as an innovative tool ensuring transparency and platform accountability, proving efficient given the monthly reports platforms submitted to the entity. The COVID-19 crisis further strengthened the Code of Practice on Disinformation (Code of Practice on Disinformation) on June 16, 2022, with co-regulation practices for major platforms. This resulted in 34 signatories, 44 commitments, and 127 specific measures related to demonetizing disinformation, enhancing political advertising transparency, expanding user tools to detect and flag false or misleading content, and establishing a robust monitoring and reporting framework at the EU and Member State levels.

Acknowledging the importance of futureproofing the Code, signatories like Facebook, Microsoft, Twitter, and YouTube agreed to expand collaboration through a permanent working group. The Code included a reinforced monitoring framework based on qualitative reporting elements and service-level indicators measuring its implementation's effectiveness. Signatories created a Transparency Center, giving the public a clear view of policies implemented to meet commitments, regularly updating relevant data. Implementing this moderation-oriented Code provided a regulatory gap coverage possibility in the U.S. American jurist Ioanna Tourkochoriti (2023) viewed the measure as an example of the "Brussels effect," signaling the EU's role as a global Internet regulator.

The Digital Services Act (DSA), a step forward in platform governance, updated rules governing digital services offered by intermediaries and online platforms in a comprehensive legal document with 155 recitals and 93 articles. Covered digital services include a wide range of online services, from simple websites to internet infrastructure services and digital platforms, online marketplaces, social networks, content-sharing platforms, app stores, and online travel and accommodation platforms.

The DSA aims to prevent illegal and harmful activities and disinformation spread, ensuring user safety, protecting fundamental rights, and creating a fair and open platform environment through clear transparency and accountability frameworks for digital platforms. User, platform, and public authority roles are rebalanced according to European values, placing citizens at the center of concerns (European Commission, FAQ, 2023).

The Digital Services Act applies directly in all EU Member States, covering online intermediaries established outside the continent offering services targeting that market. Those not established in the EU must appoint a legal representative, as many companies already do under other legal instruments. Intermediary services also benefit from clear legal liability exemptions and a single set of rules when providing services to the EU (art. 13).

To provide a harmonized framework regarding information obligations and promote a safe, predictable, and reliable environment respecting fundamental rights, the regulation distinguishes different intermediary types based on service size and nature: a) Network infrastructure services like internet access providers and domain name registrars; b) Hosting services like cloud services and web hosting; c) Online platforms connecting sellers and consumers like online marketplaces, app stores, collaborative economy platforms, and social media platforms; d) Very large online platforms (VLOPs) and search engines (VLOSEs) representing specific risks in spreading illegal content and social harms, with more than 45 million monthly users in the EU.

In general, all platforms, except the smallest (employing fewer than 50 people with annual turnover and/or total annual balance sheet not exceeding €10 million), must establish complaint and redress mechanisms and extrajudicial dispute resolution mechanisms, cooperate with trusted flaggers, take measures against abusive notices, handle complaints, disclose third-party supplier credentials, and provide online advertising transparency. The DSA standards for intermediary services provision in the internal market include, per Article 1(2): a) A regime for conditional liability exemption for intermediary service providers; b) Due diligence obligations for specific categories of intermediary service providers, focusing on transparency rules and establishing a digital service provider and cyberspace application supervision regime; c) Application and enforcement rules, including cooperation and coordination between competent authorities.

Audit is a central issue. The EU issued Delegated Regulation (EU 2024/436 of the Commission of October 20, 2023), setting norms for independent audits, especially regarding legal obligations and procedural steps ensuring task independence, absence of conflicts of interest, expertise, and professional ethics. The regulation details methodology to prevent audit organizations from issuing incorrect opinions or reaching flawed compliance conclusions.

Audit service providers must consider three risk types: (i) Detection risks—when the audit organization fails to identify relevant distortions for compliance assessment by the platform; (ii) Inherent risks—related to non-compliance intrinsic to the service's nature, design, activity, and use, as well as its provision context; (iii) Control risk—occurs when the provider's internal controls fail to prevent, detect, and correct significant distortions in advance (European Union, 2023).

Fully effective since February 17, 2024, the DSA requires all platforms and intermediary services to publish annual content moderation audit reports, including information on judicial or administrative authority orders for content removal, content removed, and automated moderation system accuracy and error rates. Virtual hosting providers must report notifications received from users and trusted flaggers, while digital platforms must report on extrajudicial dispute resolution and user suspensions for service misuse (DSA, art. 15).

In summary, the new rules create mechanisms for the European Commission and Member States to coordinate actions ensuring appropriate implementation of intermediary service obligations, including: a) Measures to combat illegal goods, services, or content online, such as a mechanism for users to flag such content and platforms to cooperate with trusted flaggers; b) New user business traceability obligations in digital markets to help identify illegal goods sellers through advanced technological solutions; c) Safeguards for users, including the possibility of challenging platform content moderation decisions; d) Obligations for VLOPs and VLOSEs to prevent system misuse, adopting risk-based measures and independent audits for risk management; e) Greater platform transparency and accountability, providing clear content moderation or algorithm usage information for content recommendations; f) Prohibition of deceptive practices and certain types of targeted advertising, such as those aimed at children and ads based on sensitive data or special categories of personal data like ethnicity, political views, or sexual orientation. Deceptive practices aiming to manipulate user choices are also banned; g) Increased transparency for all digital platform advertising and influencer communications; h) Ensuring high levels of privacy and security for minors, considering that targeted advertising based on children's profiles is no longer allowed; i) Allowing researchers access to major platform data to thoroughly examine their functioning; j) Codes of conduct and technical standards to help platforms and other stakeholders comply with legal rules. Other codes may be issued with measures to ensure platform accessibility for people with disabilities, for example.

DSA and Freedom of Expression

The European Union advocates that the DSA is a tool protecting freedom of expression, including media freedom and pluralism. It achieves this by carefully balancing rules to combat illegal content with safeguards for freedom of expression and information online. This includes protection against government interference in people's freedom of expression and information. The horizontal rules against illegal content are carefully calibrated and accompanied by robust safeguards to avoid over-removal and under-removal of content for illegality reasons.

The content moderation rules in the DSA include essential safeguards protecting the right to freedom of expression: a) The DSA requires platforms to be transparent in their content moderation. When an account is restricted, the user must be informed and has the right to appeal the decision. Service providers must also disclose their content moderation policies and how they are implemented, building trust and clear communication between platforms and users. For very large platforms, users can better understand how they affect societies, and thus, they must mitigate these risks, including regarding freedom of expression; b) Platforms must have an internal complaint handling system. Article 20 of the DSA mandates platforms provide an accessible online system functioning as a sort of dual appeal level. For at least six months, users can challenge any decisions made by the

company after receiving a notification or voluntary content moderation that: (i) removes, blocks, or restricts content visibility; (ii) suspends service provision to users; (iii) suspends or terminates accounts; or (iv) ceases or restricts the ability to monetize user-provided information. The system must also cover decisions deeming content legal or compatible, thus keeping it on the platform. Appeals must be handled attentively, non-discriminatorily, diligently, and non-arbitrarily by platforms and, if justified, must revoke the initial decision "without undue delay"; c) Platforms must provide mechanisms for users to appeal moderation decisions related to content removal, including when based on platform terms and conditions. If an account or content is suspended or limited, users have the right to contest the decision, ensuring non-arbitrary decisions and empowering users to protect their online presence. Users can appeal directly on the platform, choose an extrajudicial dispute resolution body, or seek redress in courts; d) The DSA introduces new tools to evaluate and correct biases in recommendation systems, aiming to create a fairer and more representative online experience, respecting user diversity and individuality; e) Under certain conditions, platforms may temporarily suspend abusive user activities. Such measures do not affect the freedom of platforms to set stricter measures in their terms and conditions for illegal content related to serious crimes, like child exploitation material. For transparency, this possibility must be indicated clearly and in detail in terms and conditions. Decisions must always be appealable and subject to oversight by the competent digital services coordinator. Platforms must issue a prior warning before suspension, including reasons and means of appeal against the decision.

Another important mechanism is trusted flaggers whose notifications are processed and decided with priority, regardless of the subject. The trusted flagger status is granted by the Digital Services Coordinator of the Member State where the applicant is established and must be recognized by all online platform providers. This status should only be granted to entities, not individuals, demonstrating specialized knowledge and skills to combat illegal content, working diligently, rigorously, and objectively. Trusted flaggers must show independence from any platform, expertise and competence in detecting, identifying, and notifying illegal content, and perform diligent, precise, and objective notifications (art. 22, paragraph 3).

Regarding terrorist content, trusted flaggers may be public entities like national law enforcement authorities or the European Union Agency for Law Enforcement Cooperation (Europol) or non-governmental and private or semi-public entities, such as organizations in the INHOPE network acting against the dissemination of child exploitation material and illegal racist and xenophobic online content. To maintain this mechanism's added value, the total number of trusted flagger statuses should be limited.

For transparency, trusted flaggers must publish detailed, easily understandable reports showing the number of notifications by virtual hosting service providers, content types notified, and measures taken by the service provider. These reports must be publicly available but without personal data.

Importantly, DSA application and monitoring depend on a shared responsibility framework between the European Commission and Member States. Generally, the Commission oversees VLOPs within the EU, while Digital Services Coordinators (DSCs) and other national regulators supervise platforms in their countries. In cases of Member State overlap, the European Digital Services Board may play an active role, detecting the need for joint investigations involving two or three Digital Services Coordinators. For very large online platforms and search engines, the Commission oversees supervision. This ensures platforms do not escape DSA application by choosing jurisdictions with lax enforcement. This framework also ensures more consistent rule enforcement and potentially addresses resource or expertise gaps among national supervisory authorities. The European Digital Services Board could become a valuable forum for exchanging ideas, driving secondary legislation creation, and deliberative verification of regulatory decisions, especially if building strong relationships with research and civil society organizations (Jaursch, 2024).

Finally, another central instrument in DSA enforcement is systemic risk mitigation (Article 35(3)). Systemic risks include disinformation, hate speech, illegal speech impacting civic discourse, electoral processes, and public safety, algorithmic manipulation, risks associated with children's exposure to inappropriate or dangerous content, and their personal data protection. Once identified, preventive measures can be set to ultimately strengthen regulation and oversight. For example, Article 34 imposes a duty on providers to conduct annual systemic risk assessments from their service design or operation, and related systems.

Risk assessment must consider search tools and large platform service mechanisms influencing systemic risks: algorithmic systems, content moderation mechanisms, terms and conditions and their application, advertising systems, and data practices. It must also consider specific regional or linguistic aspects, observing Member State diversity. The European Commission can identify specific systemic risks and set platform action rules. For example, on March 26, 2024, preceding the European Parliament elections, the Commission advised VLOP and VLOSE providers to strengthen internal processes to adapt mitigation

measures to identified risks in electoral processes, using available risk analysis and local context information on user service usage before, during, and after elections. Mitigation measures extended to clearly labeling generative AI content as deepfakes.

Online Safety Act 2023

In the UK, the Online Safety Act (OSA), passed in October 2023, aims to regulate online services to promote a safe environment. The law assigns duties and responsibilities to search services and web providers to minimize children's and young people's exposure to illegal or harmful content and activities. The law is effective, with implementation phases supervised by the UK's communications regulator, the Office of Communication (Ofcom).

According to the UK government (2023), "the law assigns legal responsibility to tech companies to prevent and quickly remove illegal content, such as terrorism and revenge pornography." Harmful content, such as bullying or promoting self-harm and eating disorders, must be prevented from being accessed by children and young people, while empowering adults and guardians with more alternatives about what they see online.

To empower adults and guardians, the law offers mechanisms to help identify illegal online content. According to the UK government (2023), the Online Safety Act offers three layers of user protection: (i) ensuring illegal content is removed; (ii) offering the option to filter unwanted content; (iii) ensuring online service providers comply with user-agreed terms and conditions.

In a post on the Ofcom website (2023), it is disclosed that the commitment to providing tools to empower adults and guardians will be part of the third phase of the plan, the last to be implemented by the regulatory body. For this phase, categorized as "transparency, user empowerment, and other obligations in categorized services," a consultation on the Ofcom's guidance draft will be released in mid-2024.

Besides minimizing access and propagation of hate speech and cyberbullying, the law also aims to combat disinformation and the spread of false and/or misleading information. In chapter 7, article 152, the Online Safety Act 2023 proposes creating a disinformation advisory committee, one of whose functions is to advise how regulated service providers should handle disinformation on their web applications. This aligns with media and information literacy practices, which, according to Carolyn Wilson et al. (2013), relate to accessing information and critically and ethically evaluating received content.

As disclosed on the UK Safer Internet Centre blog (2023), the law's approval was celebrated by various organizations and prominent figures who campaigned for years to raise awareness of the need for online safety regulation, especially for children and teenagers. Key factors leading to the law's creation include growing internet use for radicalization, personal data misuse by digital platforms (as demonstrated by the Cambridge Analytica case in 2018), increasing public pressure to combat hate speech, and the COVID-19 pandemic, which expanded online tool availability. Being an old demand, regulations involving safer internet norms and guidelines did not start with the Online Safety Act.

Before this new law, the *Online Harms White Paper* was a 2019 UK government report proposing regulation to combat harmful content spread on the internet, suggesting an independent regulator to draft codes of conduct for tech companies, outlining their new statutory "duty of care" to users. The document underwent a 12-week consultation period for organizations, experts, businesses, and the public to contribute to the debate. Each antecedent factor, public consultation, and published report led to the final version of the Online Safety Act.

The OSA applies to online services, whether websites or apps. The rules cover illegal online services and activities, including child sexual exploitation, fraud, and terrorism. The law introduces new online safety regulations and platform accountability, also applied to content children are exposed to, such as content promoting bullying, self-harm, and eating disorders. Regarding online service provision, the OSA applies in three specific cases: (i) sites where users can create and share content and/or interact with others, like social networks, online games, cloud file storage services; (ii) search sites; and (iii) sites publishing and displaying pornographic content (McCarron, 2023).

According to Ofcom (2024), the rules apply to all organizations, regardless of size and geographic location. Even if the company or online service is from another country, the OSA applies if it has a significant UK user base or targets the UK market. The exception is penalty enforcement, applied differently in the offending company's country of origin.

The OSA's central concept is the duty of care. According to Nóbrega (2023), the concept "assumes that individuals and legal entities must take necessary and reasonable measures to prevent damage from their activities." Thus, the Online Safety Act "establishes duties requiring all covered platforms to proactively manage online harm risks related to their services" (Kira & Schertel Mendes, 2023).

According to part 3 of the OSA, the duty of care relates to digital platform content and activity and search service providers. The law also imposes duties on platforms and providers to assess whether a service is likely to be accessed by children or involves fraudulent advertising. These duty of care measures focus on risk assessment, user protection, and greater transparency.

As Wallace (2023) explains, the Online Safety Act's duty of care covers three main categories: (i) "Illegal content (particularly content related to specific criminal offenses); (ii) content that is legal but harmful to children; and (iii) fraudulent advertising." Regarding the first category, the legislation distinguishes between illegal and priority illegal content.

In the illegal content category, chapter 2 assesses the risk of illegal content considering factors like user base and risk of accessing the material, content's likelihood of being used to commit crimes, nature and severity of harm, and dissemination potential. According to Kira and Schertel Mendes (2023), the Online Safety Act defines illegal content as situations where using words, images, speech, or sounds constitutes an offense, where possessing, viewing, or accessing such content is an offense, or when publishing the content itself is an offense, also covering everything already or becoming an offense after OSA approval.

Regarding priority illegal content, chapter 7 defines it as related to terrorism, child sexual abuse, or any offenses specified in "schedule 7," including suicide encouragement, death threats, public order offenses, harassment, stalking, or violence incitement, fraud, and other incipient offenses.

In the legal but harmful to children content category, part 61 of chapter 7 details: pornographic content, content providing suicide or deliberate self-harm instructions, and content providing eating disorder instructions. These materials are harmful whether text-only or accompanied by another format, like GIF or emoji.

Chapter 5 addresses "fraudulent advertising" duty of care, proposing that online service providers "prevent individuals from encountering fraudulent advertising through the service and minimize the time such content is present." The legislation suggests providers promptly remove fraudulent ads upon being alerted to their presence.

Part 10 of the law, effective January 31, 2024, addresses communication crimes. The provision includes sending unsolicited sexual communications and images in response to emerging online behaviors. New identified offenses punishable by law include: a) Sending false communications. A person is guilty of this offense if they intentionally transmit information, they know to be false. b) Sending threatening communications. This occurs when a person transmits a death threat or serious injury (severe bodily injury, rape, penetration assault, or severe financial loss), using serious financial loss threats as an appropriate means of enforcement. c) Sending or displaying electronically flashing images (cyberflashing) to people with epilepsy intending to harm them. d) Encouraging or aiding severe self-harm of another person. This offense can be committed regardless of self-harm occurrence and irrespective of who created the content. The sender is guilty of the offense. e) Intentionally sending a photograph or film of one's genitals to another person to gain sexual gratification, cause alarm, distress, or humiliation.

In summary, the law adopts a zero-tolerance approach to protecting children, ensuring social media platforms are responsible for the content they host. According to the UK government (2023), platforms must quickly remove illegal content or prevent its appearance, including content promoting self-harm, and prevent children from accessing harmful, age-inappropriate content, including pornography, content promoting, encouraging, or providing self-harm or eating disorder instructions, or representing and encouraging severe violence or bullying content. Platforms must impose age limits and use verification measures where harmful content is published for children, ensure social media platforms are transparent about risks and dangers to children on their sites, including through risk assessments, and provide clear and accessible ways for parents and children to report online problems when they arise.

Together with the law's publication, Ofcom was formalized as the online safety regulator, responsible for overseeing and enforcing duty of care compliance. The regulator plans to publish phase two proposals in spring 2024, including guidance drafts on child access and safety assessments and codes of practice for child online protection measures. Final guidelines will be published in early 2025, giving services three months to complete relevant assessments. Additionally, by spring 2025, Ofcom plans to publish a guidance draft on protecting women and children. The third phase addresses transparency, user empowerment, and other obligations in categorized platforms. This phase will be relevant for services designated as Categories

1 and 2B (user-to-user services) or 2A (search services), subject to additional duties such as transparency reports, user empowerment, fraudulent advertising, and user rights.

Comparison between OSA and DSA

The Digital Services Act (DSA) of the European Union and the UK's Online Safety Act (OSA) are two significant legislations regulating the digital space, but they approach the issue differently and have different scopes. The table below compares the two regulations in terms of their objectives, scope, main provisions, and expected impact based on analyses by Kira and Mendes (2013), Silkin (2023), and Dees (2023).

Category	Digital Services Act	Online Safety Act
Objectives	Create a safer and more responsible digital space where user rights are protected and online platforms are held accountable for hosted content. Ensure a high level of privacy, freedom of expression, transparency, and safety on the internet for EU users.	Focus on protecting users from harmful content on the internet, with special emphasis on children's online safety. Imposes duties on digital platforms to manage and mitigate risks associated with illegal and harmful content.
Scope	Applies to a wide range of digital services operating in the EU, including online platforms, social networks, data hosting services, and search engines. Also applies to companies both within and outside Europe, provided they offer services to users in the region. Exempts micro and small enterprises from some requirements, giving them more time to implement others. Focuses on transparency and user rights. DSA will have to expand its scope to cover Large Language Models (LLMs).	More focused on social media platforms, search engines, and other services allowing user interaction or content sharing. Applies mainly to companies operating in the UK, but may also affect international companies offering services to UK users. Exempts certain services like email, SMS, and MMS communication providers and education or childcare service providers. Strong focus on protection against harmful and illegal content, especially children's safety. OSA addresses new communication forms, like LLM chatbots.
Content	Covers a wide range of issues, including intellectual property (IP) violations, illegal goods, dark patterns, and crisis response. "Illegal content" refers to information illegal or related to illegal activities. Examples include sharing child exploitation images, selling products or services violating consumer protection law, and illegal animal sales. Legal but harmful content (e.g., bullying or self-harm) is not defined or subject to removal obligations.	Primarily deals with illegal content, but within this focus, it is more detailed and granular. OSA defines two illegal content categories—general crimes and "priority crimes" like child exploitation and terrorism, with detailed specifications. Harmful content for children and adults will be defined in secondary regulations.
Obligations	Strict rules for removing illegal content and platform liability. Transparency obligations for recommendation algorithms. Child online protection. User complaint and appeal mechanisms. Defines criteria for Very Large Online Platforms (VLOPs) and Very Large Online Search Engines (VLOSEs) subject to stricter obligations. Exempts micro and small enterprises from some requirements, allowing more time to implement others.	Legal duty of care for platforms to protect users from illegal and harmful content, adopting measures to prevent or mitigate reasonably foreseeable harms. Specific requirements for child online protection. Strong enforcement measures with significant penalties for non-compliance. Avoids setting explicit limits or criteria for each category.
Enforcement	Operates under a notice and takedown complaint system, usually using form-based procedures.	Platforms must assess risks associated with illegal content and take proportional measures to

Category	Digital Services Act	Online Safety Act
	Besides platforms and search engines, hosting service providers must implement notice and action mechanisms for illegal content. Online platforms must prioritize trusted flagger notifications (certified bodies with expertise in identifying and reporting illegal content). VLOPs and VLOSEs must implement reasonable, proportional, and effective risk mitigation measures for specific systemic risks identified. All service providers must report on the number of orders received from authorities, actions taken, complaints, automated content moderation use, and other content moderation information. Online platforms must implement a complaint handling system for content moderation decisions. Imposes transparency requirements for recommendation systems, allowing users to disable them. Requires service providers to designate a contact for communication with authorities and a user contact point.	mitigate these risks. Requires active content monitoring, shifting from the traditional voluntary approach, aiming to prevent users from encountering clearly illegal content. Services likely accessed by children must assess specific child harm risks and take proportional measures. Platforms must establish effective processes and systems for users to report illegal content, implement complaint and redress procedures, and promptly remove illegal content upon awareness. OSA obligations depend on platform type and size and content's illegal nature and child access likelihood. Lack of clarity on how platforms should age-restrict harmful content sparked privacy and freedom of expression debates.
Risk Assessment	Service providers must adopt measures for specific illegal content (if aware or notified), including removal or information provision if national authorities order. VLOPs and VLOSEs must produce annual systemic risk assessments from their service design, operation, and use concerning illegal content spread and use proportional measures to mitigate identified risks. They must conduct ad hoc risk assessments before adopting features significantly affecting identified risks. Online platforms must implement a complaint handling system for content moderation decisions. Users can hire a certified extrajudicial dispute resolution body to resolve content moderation and service suspension disputes.	User-to-user service providers must conduct illegal content risk assessments. They must assess child access to verify service or content access likelihood and conduct risk assessments for children. Services must notify Ofcom about harmful content presence. Category 1 services (larger, popular sites with many users) must also conduct user empowerment risk assessments. Services must take proportional measures to prevent users from encountering priority illegal content, manage and mitigate priority offense service use risk, and harm risks to individuals. Providers must use proportional systems and processes to minimize priority illegal content presence duration and promptly remove it upon awareness.
Impact	Expected to standardize online operation rules within the EU, offering users more security and control over online experiences. Businesses face more regulations but have a clear rule set to follow.	Aims to make the UK the safest place in the world to be online, especially for children. Platforms must invest significantly in online safety and content moderation mechanisms to comply with new regulations.

Sources: Kira and Mendes (2013), Silkin (2023), Dees (2023)

Brazilian Bill No. 2630/2020

Brazil has historically provided curious and innovative responses to some internet-related problems. Two standout initiatives are:

1. **Marco Civil da Internet (MCI)** - Internet Civil Act -, Law 14.965/2014, approved in April 2014. This law placed Brazil at the forefront of internet regulation debates, particularly regarding user rights and duties on the global computer network.
2. **Lei Geral de Proteção de Dados Pessoais (LGPD)** – General Law of Data Protection -. While not solely focused on the digital world, it rode the wave of European debates on tech company rules ensuring privacy and data protection in the digital world.

These experiences show how Brazil, at different times and political contexts, approved regulations addressing key aspects of internet regulation.

The latest regulatory debate on Brazil's digital environment agenda is Bill 2630/2020, popularly known as the "Fake News Bill." Senator Alessandro Vieira (Cidadania/SE) introduced the bill in May 2020. It proposes the Brazilian Law of Freedom, Responsibility, and Transparency on the Internet.

This legislative proposal emerged after a contentious 2018 election, where disinformative content and its proliferation through applications used by millions of Brazilians, like WhatsApp and social networks, characterized the electoral process. Vieira's bill aimed to establish norms, guidelines, and transparency mechanisms for social networks and private messaging services via the internet, discouraging their abuse or manipulation, potentially causing individual or collective harm. Applications subject to penalties under the senator's bill would be those offering social network services to the Brazilian public with over two million registered users. Applications with fewer registered users would use the proposal as a parameter for best practice programs, adopting adequate and proportional measures to combat disinformation and ensure transparency about paid content. All applications, including those offered by foreign legal entities, must respect the legislative proposal's rules if they offer services to the Brazilian public or if at least one group member has an establishment in Brazil.

Senator Alessandro Vieira sought to impose rules to mitigate problems arising from Brazilian elections. His version of the bill aimed to strengthen the democratic process by combating disinformation and fostering information diversity on the internet in Brazil, increasing transparency about paid content available to users, and discouraging the use of inauthentic accounts to disseminate disinformation.

In the Senate, at least five reports were presented by the proposal's rapporteur, Senator Angelo Coronel (PSD/BA). In June 2020, under various interest group pressures, the Senate approved a version different from Vieira's. The approved text by Angelo Coronel included changes such as requiring operators to send customer data to messaging services (e.g., WhatsApp, Telegram) of customers who had their contracts terminated. The previous text required operators to send data of disabled numbers, not terminated contracts. Messaging platforms exclusively linked to cell numbers could suspend accounts with terminated contracts using information provided by operators. The text clarified that only phone numbers would be provided, without additional customer data.

The Senate's approval was seen as rapid, and after various sectors' pressures, the final text became less punitive, creating transparency procedures for internet application content management practices. However, the bill tended towards becoming a monitoring instrument, creating vulnerabilities in user data collection.

In the House of Representatives, then-House Speaker Rodrigo Maia determined the proposal be analyzed by a working group (WG). The WG's rapporteur was Congressman Orlando Silva (PCdoB-SP), with Congresswoman Bruna Furlan (PSDB-SP) chairing. In the House, the text needed more discussion, indicating longer debates than in the Senate. Thus, the WG, chaired by Bruna Furlan and reported by Orlando Silva, held 15 public hearings to hear various sectors on the issue. Despite the Senate's approval, interest groups felt the topic needed more debate.

Interest group disputes—broadcasting companies, big techs, civil society, and telecom operators—intensified in the House. Google's president, Fabio Coelho, published a letter stating that Bill 2630/2020, as it stood, would facilitate disinformation dissemination rather than combat it.

The matured WG text imposed obligations making search tools less secure and more susceptible to abuse and fraud. The main obligation was detailed information disclosure on system functioning, part of requiring search platform transparency. The broadcasting sector, organized in the Freedom with Responsibility Coalition comprising 43 national and state entities of communication companies and professionals, supported the Fake News Bill, requesting House Speaker Arthur Lira (PP/AL)

and Senate President Rodrigo Pacheco (PSD/MG) prioritize the matter. Coalition members include broadcasting, journalism, advertising, and marketing entities.

The final WG text presented by Orlando Silva aimed to: a) Prohibit creating fake social media accounts simulating a person or entity's identity; b) Ban bot use—automated accounts managed by robots; c) Limit highly shared message reach; d) Require companies to keep records of mass-forwarded messages for three months; e) Require user identification for sponsored content to prevent false financial scam ads; f) Prohibit official government or public interest organization accounts (e.g., politicians) from blocking citizen accounts; g) Create the Internet Transparency and Responsibility Council, an autonomous supervisory entity to regulate and oversee providers; h) Require social media providers to establish headquarters in Brazil; i) Impose sanctions or penalties, such as warnings or fines, on companies violating bill measures if it becomes law.

To facilitate approval, the bill excluded journalistic content remuneration provisions. This topic migrated to Bill 2370/2019 by Congresswoman Jandira Feghali (PCdoB-RJ), focusing on comprehensive internet copyright reform. News content remuneration aims to compensate journalism companies for content used by providers. Thus, Brazilian communication sector entities and journalism organizations supported the Fake News Bill.

Despite this exclusion, the bill's approval was delayed. In May 2023, it was on the House's plenary agenda but removed at Orlando Silva's request. He sought more time to adapt the report to new contributions from various groups.

The bill's main issue was the supervisory entity for platform obligations and responsibilities. Rapporteur Orlando Silva noted this aspect still needed resolution. Sectors wanted the National Telecommunications Agency (Anatel) or a tripartite system proposed by the OAB Digital Rights Special Committee (CEDD).

The OAB Digital Rights Special Committee proposed a Tripartite Brazilian Platform Regulation System, including user, regulatory agency, civil society, academia, business, and executive, legislative, and judicial representatives to ensure a checks-and-balances system. The OAB advocated judiciary, communication regulatory agency, House, and Senate representatives.

To contribute to the supervisory body debate, the executive branch suggested creating an autonomous supervisory entity. Without much detail, Digital Policy Secretary João Brant explained it would supervise platform self-regulation, not individual content. Without a defined structure, the format would be discussed with society for the best government path.

In the bill's latest report, Orlando Silva removed the executive branch's autonomous platform supervision entity creation provision, attempting bill approval. However, the proposed text requires at least 26 further regulations without specifying the responsible entity or deadlines. Without these regulations, most mechanisms in the bill cannot be applied.

The Brazilian regulation proposal has similarities with the DSA regarding platform obligations: a) Systemic risk analysis and mitigation; b) Independent audit requirement, though questioned by legal scholars due to potential "audit washing," where companies set standards for a false compliance guarantee; c) Platform data access for researchers; d) Transparency reports with clear content moderation activity information; e) Crisis response protocols for extraordinary circumstances posing serious public safety or health threats.

In other aspects, the Brazilian project differs from the DSA, such as the duty of care in the proposal, more akin to Germany's NetzDG law, which has a list of illegal content for removal by platforms without a court order. Brazil still discusses the supervisory model and entity for regulation enforcement. A consensus is difficult for a structure like the European Digital Services Committee. Another difference is social media use by state representatives, known as parliamentary immunity. Content moderation rules gain importance as they can affect specific individuals and groups' freedom of expression.

However, the ongoing debate highlights the importance of balancing market, digital rights, and freedom of expression in regulations. They may be segmented into specific laws but must remain interconnected.

Conclusions

The Digital Services Act (DSA) addresses the European continent's need for a transparency and accountability standard for major platform content moderation and algorithm use.

It is a comprehensive internet user protection regulation seeking to balance freedom of expression protection and hate speech limitation. It maintains fundamental principles of the E-Commerce Directive, like conditional hosting provider immunity and general monitoring prohibition. Transparency obligations for intermediaries and large online platforms include detailed reports on content moderation, automated means use, and content removal decision justification, maintaining complaint and appeal mechanisms for users. It uses a good-faith responsibility model. Providers and platforms proactively removing illegal content are immune from liability unless they actively disseminate or promote illegal content beyond technical processing.

Despite this concern, a challenging aspect for freedom of expression is the broad definition of "illegal content" in the law. Illegal content includes information in any form, such as hate speech, terrorist content, and illegal discriminatory content. The broad concept allows platform and authority discretion, possibly leading to excessive content removal affecting users' freedom of expression.

Partially, the lack of detail on what constitutes illegal content may be offset by systemic risk mitigation mechanisms. In the DSA context, systemic risks arise from platform design and operation with potential or actual negative societal or fundamental rights impacts. Systemic risks include content with negative effects on fundamental rights, including human dignity, privacy, freedom of expression, non-discrimination, children's rights, and consumer protection, impacts on electoral processes, public safety, gender violence, public health, minors, and individuals' physical and mental well-being. The DSA requires VLOPs to assess and take reasonable, proportional, and effective measures to mitigate these systemic risks, including content moderation system adjustments, content recommendation system changes, or crisis protocols. The European Commission can issue systemic risk guidelines considering specific regional or linguistic aspects and Member State diversity.

The DSA's systemic risk approach incorporates flexibility and adaptability to new knowledge, socioeconomic processes, and technological developments. However, its effectiveness depends on shared implementation and enforcement by the European Commission, European Digital Services Committee, Member States, and platforms. A critical point is the broad discretion given to national authorities to order content removal. For example, the DSA allows national authorities to issue orders against specific illegal content based on Union or national legislation consistent with Union law. Excessive content removal could negatively affect users' freedom of expression.

There are also concerns about the European Commission's supervisory powers, especially for VLOPs, ensuring DSA obligations compliance. The non-governmental organization Article 19 argues that the Commission's DSA supervision is not entirely independent, as it is the EU's executive branch, raising concerns about balancing platform regulation and freedom of expression protection.

The UK, outside the EU, created the Online Safety Act (OSA) to balance protection against harmful and illegal online content with freedom of expression preservation. It includes specific clauses covering freedom of expression and privacy for all services, democratic content, news publishers' content, and journalistic content, especially for Category 1 services, the most significant and influential. These services have an additional duty to conduct and publish updated impact assessments of their safety measures and policies, specifying positive actions taken in response to these assessments to protect users' freedom of expression and privacy rights.

Critics argue that subjective censorship standards in the regulation allow the UK government to silence users and imprison those posting disapproved messages. Section 10 of the OSA authorizes up to two years in prison for social media messages causing "psychological harm equivalent to at least severe distress." The law is also seen as extremely complex, potentially undermining freedom of expression, information, and privacy rights. Platforms must enforce their terms of service, potentially leading to overzealousness and excessive censorship power.

Conversely, the OSA better defines illegal content, being more detailed and granular. It outlines two illegal content categories—general crimes and "priority crimes" like child exploitation and terrorism. Harmful content for children and adults will be defined in secondary regulations. Like the DSA, the UK law adopts a risk assessment mechanism for illegal content, requiring platforms to assess child access to verify service or content access likelihood and conduct risk assessments. The regulator, OFCOM, must be notified of harmful content presence. Services must take proportional measures to prevent users from encountering priority illegal content, manage and mitigate priority offense service use risk, and harm risks to individuals.

The Online Safety Act is an attempt to make the online environment safer and protect user rights, raising questions about balancing safety and freedom of expression. An issue is the Secretary of State's excessive power in directing the media regulator

(Ofcom), as the law allows altering the priority illegal content list, such as terrorism, requiring stricter platform actions, giving high control over online content removal.

Similarly, the solution for disinformation and sensitive content suggests "filters" for platforms, allowing adults to choose what they see or don't, not necessarily solving the problem as platforms need only offer non-display options. Diverse groups fear that in trying to protect children and make the internet safer, the law may censor certain content types and deprive users of freedom of expression.

While Europe advances in platform regulation, Brazil still discusses Bill 2630/2020, known as the Fake News Bill. The debate often polarizes between those seeing it as a means to combat disinformation and protect online human rights and those fearing freedom of expression and privacy restrictions.

The bill proposes measures to prevent, combat, and hold accountable false news dissemination on digital platforms, inspired by the European Digital Services Act. It suggests creating an autonomous entity to oversee internet providers, called the Internet Transparency and Responsibility Council, to initiate administrative proceedings and apply sanctions for non-compliance. Platforms must act quickly to prevent and combat illegal practices, being civilly liable for illicit content circulation and informing authorities about any suspected crime information. Additionally, users must be notified about content removal or concealment reasons, with the possibility of contesting decisions.

However, this law may conflict with Brazil's constitutional freedom of expression principle, established in Article 5, paragraphs IV and IX, guaranteeing freedom of thought and expression, regardless of censorship or license.

Balancing market, digital rights, and freedom of expression in regulations is a significant challenge for a polarized Congress. The common point among the three regulatory frameworks is lacking mechanisms to change business models, inhibit or eliminate personal data use for targeted advertising, especially politically motivated.

Platform remuneration systems for users are fundamental in circulating and promoting illegal content and disinformation. Platforms generally have service terms and policies prohibiting such practices, suspending monetization for violators. However, monitoring mechanisms to detect policy violations are often automated or through non-assessable decision-making processes. The volatile environment makes content moderation alone insufficient, as influencers may migrate to platforms with permissive service terms, and extremists adapt to moderation using evasion techniques. This indicates a need for flexible regulatory frameworks, capable of adapting to daily changes, and a holistic approach to addressing illegal content promotion and disinformation, focusing not only on content but also on the processes and platforms producing it.

Finally, Jürgen Habermas, in his profound and erudite archaeology of the public sphere category and its structural change in the capitalist stabilization process of bourgeois society after the 17th and 18th-century revolutionary cycles, demonstrated how choosing a business model in the transforming press industry can unpredictably impact the social fabric.

If Benjamin Day foresaw only business success in offering the *New York Sun* for 1 penny in exchange for massive commercial advertising, William Randolph Hearst and Joseph Pulitzer combined news practice and profit with personal and political goals, even at the cost of pushing truth-seeking to unpredictable limits. The emotion blurring reason, even at the cost of "fostering" a war, was the key to commercial success, placing them among the wealthiest Americans of their time, risking ethical and normative principles to which John Milton, Thomas Paine, and John Stuart Mill dedicated precious years.

Perhaps that's why, when the printed, journalistic, and cultural information universe met the electronic information universe in the form of radio and television, Europe, unlike the United States, chose a non-commercial business model, taking advantage of the new journalistic and cultural information industry needing state authorization to use the common good of radiofrequency spectrum. Even in the US, always faithful to radical free markets, the new industry was regulated by an autonomous state entity, initially technical but soon addressing content issues, from obscenity to ensuring equal time for opposing election candidates.

The internet was born state-owned, as a defense infrastructure tool during the Cold War. When this seemed definitively over (which it wasn't), it remained state-owned, given to a science and research organization that soon released it for commercial use. Tim Berners-Lee, when seeing the potential of a global information exchange network between individuals and institutions, worked for a public, supranational research entity, the European Organization for Nuclear Research (CERN), in Switzerland. When he transformed his vision into the World Wide Web protocol system in 1989, the rights were publicly held by CERN.

However, once commercially liberated, the Web quickly became the political-economic force leading a new capitalism stage. As pointed out in a *Fortune* magazine report on July 10, 1995, the Web already indicated a business model preference for one side of US politics and potential compromises for culture and democracy. These movements gained uncontrollable traction when Google, now Alphabet, launched its renowned search tool, commercialized by massively using user data, driving a sophisticated programmatic advertising system.

Paradoxically, the company born under the "Do no evil" motto, later changed to the more realistic "Do the right thing," is at the origin of the structural, existential crisis plaguing "big techs"—Google, Facebook (now Meta), Instagram, WhatsApp, Amazon, and others. This crisis centers on this study's subject: freedom of expression in the digital ecosystem, threatened by disinformation and attempts to counter it, potentially jeopardizing freedom of expression on the Web.

This study, within its scope, demonstrated through analyses of the Digital Services Act, Online Safety Act, and Brazilian Bill No. 2630/2022, that it's possible, within acceptable regulatory limits, to inhibit Web disinformation. However, both normative instruments still need efficacy testing.¹

Therefore, in the broader context of the New Regulatory Challenges of the Digital Ecosystem research, the study advanced towards more significant normative and regulatory impact on Web information services, addressing not only proven harmful effects but the business model itself. This led the research back to Tim Berners-Lee's Solid project at MIT's Computer Science and Artificial Intelligence Laboratory, a decentralized information group acronym for Social Linked Data. Solid's technical specifics and ethical potential, including business developments since 2017 through the Inrupt startup, are part of the research's final report. Notably, Solid is being used in a broad range of public policies in Belgium.

For now, it's enough to say that Solid attempts to challenge the current business model based on massive personal data exploitation through a protocol and an application, a Personal Data Storage (POD), where user data is securely stored with varying access levels. Users can then consent and share their data with whom they wish and for desired purposes, with data returning to their POD after consent expiration.

Ultimately, Solid does not harken back to the legacy that currently defines platforms, not only those centered in the information market, but all of those that control what Berners-Lee calls "silos," where the data they collect, process, and profit from reside massively. Nonetheless, from the perspective of this paper, especially in light of its revelations about the limitations that current European and British legislation attempt to impose on the spread of misinformation, the idea put into practice around Solid only confirms that, aside from a structural change in their current business models, legislation will continue to follow in an eternal game of catch-up, weakening freedom without fundamentally addressing the core problem.

REFERENCES

1. Adorno, Theodore and Horkheimer, Max (1985). *Dialética do Esclarecimento*. Rio de Janeiro, Zahar.
2. Artigo 19 (2021). *At a glance: Does the EU Digital Services Act protect freedom of expression?*. Available in: <https://www.article19.org/resources/does-the-digital-services-act-protect-freedom-of-expression/>
3. BBC (2018). Reino Unido aprova leis que podem punir até com prisão criadores de 'deepfakes'. *BBC News*. Disponível em: <https://www.bbc.com/portuguese/internacional-43466255>.
4. Brasil (2020). Senado Federal. Projeto de Lei 2.630, de 13 de maio de 2020. "Institui a Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet". Brasília: *Senado Federal*, 2020. Available in: <https://legis.senado.leg.br/sdleg-getter/documento?dm=8110634&ts=1697571574911&disposition=inline>
5. Comissão Europeia (2023). *Digital Services Act*: Commission designates first set of Very Large Online Platforms and Search Engines. Press Release, 25.04. Available in: https://ec.europa.eu/commission/presscorner/detail/en/IP_23_2413
6. Comissão Europeia (2023). *DSA: Very large online platforms and search engines*. Shaping Europe's digital future. Available in: <https://digital-strategy.ec.europa.eu/en/policies/dsa-vlops>
7. Comissão Europeia (2021). *EU Digital Markets Act and Digital Services Act explained*. Topics 14.12. Disponível em: <https://www.europarl.europa.eu/topics/en/article/20211209STO19124/eu-digital-markets-act-and-digital-services-act-explained>

¹ The reports of the studies carried out for this project are available at: <https://www.gov.br/anatel/pt-br/centrais-de-conteudo/publicacoes/estudos-e-pesquisas/estudos-descentralizados>

8. Comissão Europeia (2022). *The 2022 Code of Practice on Disinformation*. Shaping Europe's digital future. Available in: <https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation>
9. Comissão Europeia (2024). Communication to the Commission Approval of the content of a draft Communication from the Commission on Guidelines for providers of Very Large Online Platforms and Very Large Online Search Engines on the mitigation of systemic risks for electoral processes pursuant to the Digital Services Act (Regulation (EU) 2022/2065). Available in: https://ec.europa.eu/newsroom/repository/document/2024-13/C_2024_2121_1_EN_annexe_acte_autonome_cp_part1_v3_tpHHZgYyBGFMF8J5rE0OR1GdOis_103911.pdf
10. European Parliament. (2022). *Digital Services Act: Questions and Answers*. Available in: https://ec.europa.eu/commission/presscorner/detail/en/qanda_20_2348
11. Felipe, Mathias. Regulações da internet pelo mundo: Reino Unido promete ser pioneiro. *Desinformante*. Available in: <https://desinformante.com.br/regulacoes-da-internet-pelo-mundo-reino-unido-promete-ser-pioneiro/>.
12. Goodman, Emma (2019). *The Online Harms White Paper: Its approach to disinformation and the challenges of regulation*. LSE. Available in: <https://blogs.lse.ac.uk/medialse/2019/04/10/the-online-harms-white-paper-its-approach-to-disinformation-and-the-challenges-of-regulation>.
13. Governo do Reino Unido (2023). Online Safety Act. Available in: <https://www.legislation.gov.uk/ukpga/2023/50/enacted>.
14. Governo do Reino Unido. UK children and adults to be safer online as world-leading online safety laws become law. Available in: <https://www.gov.uk/government/news/uk-children-and-adults-to-be-safer-online-as-world-leading-online-safety-laws-become-law>.
15. Governo do Reino Unido (2022). World-first online safety laws introduced in Parliament. Available in: <https://www.gov.uk/government/news/world-first-online-safety-laws-introduced-in-parliament>.
16. Greenblatt, Stephen (2018). *Ascensão e Queda de Adão e Eva*. São Paulo: Companhia das Letras.
17. Habermas, Jürgen (1984). *Mudança Estrutural da Esfera Pública – Investigação quanto a uma categoria da sociedade burguesa*. Rio de Janeiro: Tempo Universitário.
18. Jaurisch, Julian (2024). More than an advisory group: why The European Board for Digital Services has key roles in DSA enforcement. *DSA Observatory*. Available in: <https://dsa-observatory.eu/2024/02/23/more-than-an-advisory-group-why-the-european-board-for-digital-services-has-key-roles-in-dsa-enforcement/#:~:text=The%20Board%20is%20an%20independent,roles%2C%20see%20Jaurisch%202024>.
19. Kira, B.; Mendes, L (2023). A Primer on the UK Online Safety Act: Key aspects of the new law and its road to implementation. *Verfassungs blog*, 13.11. Available in: <https://verfassungsblog.de/a-primer-on-the-uk-online-safety-act/>.
20. McCarron, Victoria (2023). Online Safety Act: who does it apply to?. Who does it apply to?. *Burges Salomon*. Available in: <https://www.burges-salmon.com/news-and-insight/legal-updates/technology-and-communications/online-safety-act-who-does-it-apply-to>.
21. Mill, John Stuart (2011). *Sobre a Liberdade*. Rio de Janeiro: Nova Fronteira.
22. Milton, John (2021). *Areopagitica – Um discurso do Sr. John Milton pela liberdade de expressão não licenciada, para o Parlamento da Inglaterra*. São Paulo: Convivium.
23. Milton, John (2016). *Paraíso Perdido*. São Paulo: Editora 34.
24. Nóbrega, Liz (2023). Reino Unido está a um passo da lei de “tolerância zero” para proteção das crianças nas redes sociais. *Desinformante*. Available in: <https://desinformante.com.br/reino-unido-protecao-criancas/>.
25. Norton Rose Fulbright (2023). Online Safety Act. *Norton Rose Fulbright*, dezembro. Available in: <https://www.nortonrosefulbright.com/pt-br/knowledge/publications/0b658a5a/online-safety-act>.
26. Ofcom (2023). Roadmap to regulation: *Ofcom's approach to implementing the Online Safety Act*. Available in: <https://www.ofcom.org.uk/online-safety/information-for-industry/roadmap-to-regulation>.
27. Ofcom (2023). *Online safety: Ofcom role and what it means for you*. Available in: <https://www.ofcom.org.uk/news-centre/2023/online-safety-ofcom-role-and-what-it-means-for-you>.
28. Safer Internet (2021). Online Safety Bill becomes law. Safer Internet. Available in: <https://saferinternet.org.uk/blog/online-safety-bill-becomes-law>.
29. Seelaender, Airton C.L (1991). Surgimento da idéia de liberdades essenciais relativas à informação- a 'Areopagitica' de Milton". *Revista da Faculdade de Direito (USP)*, S. Paulo, v. 86.
30. The Guardian (2023). How the UK's online safety bill aims to clean up the internet. TechScape, outubro. Available in: <https://www.theguardian.com/technology/2023/oct/24/techscape-uk-online-safety-bill-clean-up-internet>.

31. Tourkochorit, Ioanna. The Digital Services Act and the EU as the Global Regulator of the Internet. *The Internet's Global Regulator*, Vol. 24 No. 1, 2023. Available in:
<https://cjl.uchicago.edu/sites/default/files/2023-07/24ChiJIntlL129.pdf>
32. Unesco (2019). *Information and Communication Technology (ICT) in Education in Asia: A Comparative Analysis of ICT Integration and e-Readiness in Schools across Asia*. Paris: UNESCO. Available in:
<https://unesdoc.unesco.org/ark:/48223/pf0000220418>.
33. Wallace, Eva (2023). A Lei de Segurança Online 2023. *Vardag*. Available in:
<<https://vardags.com/posts/online-safety-act-2023>>.
34. Wilson, Carolyn; Grizzle, Alton; Tuazon, Ramon; Akyempong, Kwame and Cheung, Chi Kim (2013). *Alfabetização midiática e informacional: currículo para formação de professores*. Unesco, 2013. Available in:
<https://unesdoc.unesco.org/ark:/48223/pf0000220418>
35. Wired (2023). The UK's Controversial Online Safety Act Is Now Law. 26.10. Available in:
<https://www.wired.com/story/the-uks-controversial-online-safety-act-is-now-law/>.
36. União Europeia (2022). *Regulamento (UE) 2022/2065* do Parlamento Europeu e do Conselho, de 19 de outubro – Regulamento dos Serviços Digitais. Available in:
<https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:32022R2065>
37. União Europeia (2023). *Regulamento Delegado (UE) 2024/436* da Comissão de 20 de outubro. Available in:
https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=OJ:L_202400436
38. Urupá, Marcos (2023). PL das Fake News tem votação adiada. *Teletime*, Brasília, ano 25. 02 de mai. Available in:
<<https://teletime.com.br/02/05/2023/pl-das-fake-news-tem-votacao-adiada/>>
39. Urupá, Marcos (2023). PL que regula direitos autorais e remuneração de conteúdo jornalístico ainda tem acordo incerto. *Teletime*, Brasília, ano 25. 10 de jun. Available in:
<<https://teletime.com.br/10/08/2023/pl-que-regula-direitos-autorais-e-remuneracao-de-conteudo-jornalistico-tem-acordo-incerto/>>
40. Urupá, Marcos (2023). João Brant explica a proposta do governo para regular plataformas e combater a desinformação. *Teletime*, Brasília. Ano 25, 6 de abr. Available in:
<<https://teletime.com.br/06/04/2023/joao-brant-explica-a-proposta-do-governo-para-regular-plataformas-e-combater-a-desinformacao/>>
41. Zinn, Howard (2005). *A People's History of the United States*. Harper Perennial.