

FROM THE IRON CURTAIN TO THE ALGORITHMIC CAGE: RECONCEPTUALIZING THE GUARANTEE OF FREE EXPRESSION IN THE AGE OF DIGITAL GOVERNANCE

Vicovan Alina

Ștefan cel Mare University of Suceava, Faculty of Law

Year of study: 3rd

Abstract

Historically, the fundamental guarantee of Free Expression, enshrined in instruments like Article 10 of the ECHR, primarily functioned as a shield against the brutal, direct censorship of the State, symbolically exemplified by regimes behind the Iron Curtain. However, the reality of the 21st century presents a new set of challenges that render these classical constitutional safeguards increasingly irrelevant.

This paper argues that a substantial Constitutional Lag exists between the normative promise of free speech and its application in the digital sphere, where discourse is increasingly regulated by the opaque internal policies and algorithms of private tech entities. The central claim is that this new structure constitutes an Algorithmic Cage.

The study examines two major threats: first, the effective shift of censorship power to private actors through deplatforming, which often lacks judicial transparency and due process; and second, the risk that Artificial Intelligence (AI), used in the war against disinformation, generates a significant chilling effect on legitimate and minority discourse. Furthermore, the paper demonstrates the inadequacy of the traditional State Action doctrine to protect the crucial Right to Visibility.

The essay concludes that a legal reconceptualization is necessary to extend fundamental rights responsibilities to private corporate power. The analysis highlights the European Union's Digital Services Act (DSA) as a key legislative attempt to impose algorithmic transparency and fair redress mechanisms, aiming to reduce the gap and prevent the guaranteed right to speak from turning into a mere "right to whisper into the void".

INTRODUCTION

Freedom of expression is, for centuries, the cornerstone of democracy and human development, being known as a first-rank constitutional guarantee in most international legal instruments, such as the European Convention on Human Rights. Historically, the fight for this fundamental right has been defined by the confrontation with direct state oppression. Totalitarian regimes, such as those behind the Iron Curtain, used the brute force of censorship, political police, and formal bans to suppress speech. In that legal reality, the threat was clear, and the enemy of freedom was the state itself.

However, as we have moved into the 21st century, the nature of the threat has evolved radically. Constitutional guarantees designed to protect the citizens from the state are proving

to be insufficient or even irrelevant in the face of new legal realities generated by technology. Today, public discourse is no longer governed by courts and parliaments, but by algorithms and the internal policies of giant private entities such as Meta and Google.

This paper analyses the gap (Constitutional Lag) that exists between the normative promises of the right to free expression and the reality of its application in the digital environment. The central thesis is that, in the absence of a legal reconceptualization of fundamental rights, we are faced with a new form of censorship, the Algorithmic Cage, in which visibility, access to discourse, and even truth is filtered and moderated by private powers, often with covert government support.

Therefore, the present study aims to examine: first, how the transfer of censorship power from the state to private actors (deplatforming) manifests itself; second, how artificial intelligence (AI) complicates the fight against disinformation, generating a chilling effect on legitimate discourse; and, finally, how the judicial system (European Court and National Courts) is trying, in a difficult way, to adapt the protection mechanisms (state action doctrine) to this new Digital Governance. Finally, the paper will propose solutions to reduce this gap, ensuring that the right to speak does not become just a right to whisper.

1. THEORETICAL AND HISTORICAL ANCHOR: FROM STATE OPPRESSION TO PRIVATE POWER

The fundamental guarantee of Free Expression was historically conceived as a vertical shield for the citizen against the overreaching arm of the State. This constitutional premise, central to Article 10 ECHR, reflects an era when the primary threat to the free exchange of ideas was government power.

This is starkly contrasted by the experience behind the Iron Curtain, where censorship was an act of direct, brutal state oppression. Agencies like the Securitate enforced ideological conformity through systematic punishment, creating a society based on fear and absolute self-censorship. The state maintained a complete monopoly on information, aggressively jamming foreign stations like Radio Free Europe.

The legal reality of the twenty-first century, however, is defined by the emergence of Platform Governance. The Internet, once a bastion of freedom, has become a territory dominated by a handful of digital sovereigns (Meta, X, Google). These private entities now control the infrastructure of public discourse.

If the past threat was the State's direct order, the current challenge lies in the arbitrary decisions made by these private actors, often based on opaque Terms and Conditions or the instantaneous verdicts of content moderation algorithms. The impact of their decisions on the audibility and visibility of speech is now far greater than that of any single national law.

This structural shift has created a Constitutional Lag. The traditional legal dogma (the State Action doctrine) which applies fundamental rights exclusively to the State-Citizen relationship, is inadequate when the public sphere has been privatized. A profound rethink is needed to transition from a purely vertical application of rights to a model of horizontal applicability. This is essential to impose fundamental rights standards (transparency, due process) on these dominant private corporations, which act, de facto, as "administrators of public space". Failure to do so risks rendering the guaranteed right to speak a mere "right to whisper into the void".

1.2. The Constitutional Promise and the ECHR Mandate

Freedom of expression is not an absolute right, but a guarantee that has evolved from a historical necessity: that of protecting the individual from abuse by the authorities. Freedom of expression is not merely a classical individual privilege, but a fundamental pillar of democracy, conceived to protect individuals from state abuse. At the European level, this guarantee is eloquently enshrined in Article 10 of the ECHR¹, which not only protects innocuous ideas, but also those that "shock, offend or disturb".

The guarantee is twofold: on the one hand, it imposes on the state a negative obligation not to interfere in the discourse of its citizens; on the other hand, it imposes on it a positive obligation to create a framework in which all voices can be heard. The classic mechanism for limiting this right (such as public order or national security) presupposes an intervention justified by law, necessary in a democratic and proportionate society. However, fundamentally, all these parameters were designed to judge the actions of the state.

¹ European Convention on Human Rights (ECHR), Article 10 – Freedom of expression:

1. Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers. This Article shall not prevent States from requiring the licensing of broadcasting, television or cinema enterprises.

2. The exercise of these freedoms, since it carries with it duties and responsibilities, may be subject to such formalities, conditions, restrictions or penalties as are prescribed by law and are necessary in a democratic society, in the interests of national security, territorial integrity or public safety, for the prevention of disorder or crime, for the protection of health or morals, for the protection of the reputation or rights of others, for preventing the disclosure of information received in confidence, or for maintaining the authority and impartiality of the judiciary.

At national level, the principle is mirrored: in Romania's Constitution (art. 30), freedom of expression is declared inviolable, and censorship is prohibited.² Nonetheless, both under the ECHR and Romanian law, freedom of expression is not absolute. Article 10 (2) allows for restrictions in a democratic society, such as protection of reputation, national security, public order, or public morals.

Jurisprudence has repeatedly emphasized that any restriction must be provided by law, pursue a legitimate aim, and be strictly necessary and proportionate; thus, ensuring that limitations do not undermine the core of the right. In this sense, the constitutional and conventional framework provides a solid anchor: the protection of controversial or dissident speech, the institutional prohibition of censorship, and a safeguard of pluralism.

However. These guarantees were originally designed to limit state action, not to regulate the complex reality of contemporary communications, where a few private platforms and powerful private actors control what becomes visible.

1.2. The Iron Curtain as Historical Contrast

For understanding how much the legal landscape has changed, it is important to anchor the discussion in the historical reality of rights violations. Behind the Iron Curtain, the legal reality was the exact opposite of the guarantee: freedom of expression was annihilated by the direct, brutal and unfiltered action of the state.³

This oppression was not just an extra-legal practice; it was enshrined in law. Although the Constitution formally proclaimed freedom of the press, legal mechanisms, such as the Press Law and the Decree on the Regime of Printing, expressly established state censorship. The General Directorate of Press and Printing (DGPT), directly subordinated to the Communist Party, had absolute authority to examine and approve (or prohibit) any form of public communication, from books and newspaper articles to letters and telegrams.

Under communist regimes, especially in Romania, The Securitate, the political police agency, functioned not only as a censorship body, but as an instrument of systematic terror designed to annihilate any form of dissent. Control was not limited to the elimination of

² *Constituția României* (The Constitution of Romania), Art. 30(1) – Freedom of expression: "*Freedom of expression of thought, opinions, or beliefs, and freedom of any creations whatsoever, by word, in writing, in pictures, or by any other means of communication in public, are inviolable.*"

³ This annihilation was formalized through legal instruments. For instance, the Direcția Generală a Presei și Tipăriturilor (DGPT), established in 1949, was the central institution of state censorship in Romania, responsible for controlling all published texts, films, and media content to ensure full compliance with the official communist ideology and the principles of Socialist Realism. See Liliana Corobca, *Controlul Cărții. Cenzura Literaturii în Regimul Comunist din România* (Cartea Românească, 2014) for a detailed analysis of DGPT's mechanisms.

inconvenient texts; it extended to physical and psychological punishments applied to those who manifested even the slightest deviation from the Party line, through networks of political prisons, forced residences and total surveillance. The goal was to create a society based on fear, where citizens were forced into total self-censorship and mutual denunciation in order to survive.⁴

In addition to direct repression, the regime established a complete information monopoly. All means of communication (print press, radio, television) were strictly subordinated to the Romanian Communist Party, with the sole role of conveying intense propaganda and building the leader's personality cult. At the same time, any external source of information that might contradict the official reality was aggressively blocked. Efforts to jam foreign radio stations (such as Europe Free or Voice of America) turned access to objective news into a clandestine act of resistance, thus highlighting the brutal, direct, and omnipresent nature of state censorship, in stark contrast to the subtlety and decentralization of today's digital control.

This period represents the archetype of the threat to rights: an oppressive political will that acts in plain sight, directly on the body and mind of the citizen. By contrast, the contemporary struggle is no longer against a visible censor, but against opaque, indirect, and private forces, which renders our legal instruments inadequate, generating a dramatic lag.⁵

2. THE ALGORITHMIC CAGE: NEW THREATS TO FREE EXPRESSION

If in the past, the threat to free speech came in the clear, unmistakable form of government order or direct state censorship (a scenario exemplified by the repressive regimes behind the Iron Curtain) the current legal reality is overwhelmingly defined by the phenomenon of Platform Governance.

⁴ The Securitate's role extended far beyond traditional content censorship, utilizing pervasive surveillance, psychological terror, and penal institutions to enforce total ideological conformity. This system effectively led to the internalization of fear and total self-censorship—the ultimate goal of the regime's repressive machinery. The jamming of international broadcasts like *Radio Free Europe* was a key component of the state's monopoly on information, attempting to create a hermetically sealed information environment. See Dennis Deletant, *Ceausescu and the Securitate: Coercion and Dissent in Romania, 1965-1989* (C. Hurst & Co., 1995); see also Cristian Vasile, *Literatura și artele în România comunistă, 1948-1953* (Humanitas, 2010).

⁵ The legal framework explicitly mandated the media's role as a political instrument. For example, Law No. 3/1974 regarding the press stipulated in Art. 2 that the press unfolds its activity "under the leadership of the Romanian Communist Party", thus transforming mass-media into a platform for implementing the Party's general policy rather than an organ of public opinion. This contrasts sharply with the contemporary challenge of *private* (non-state) ideological control. See LEGE 3 28/03/1974 – Portal Legislativ.

The Internet, which was once hailed as a sprawling, decentralized space of absolute freedom and egalitarian discourse, has paradoxically become a territory dominated by an astonishingly small number of private, transnational companies, such as Meta (Facebook, Instagram), X (formerly Twitter), and Google (YouTube). These entities do not merely host content; they effectively control the infrastructure of public discourse on a global scale.

The decisions made by these corporate actors, which determine the reach and survival of ideas, are often based on opaque Terms and Conditions or the instantaneous, unappealable verdicts of content moderation algorithms. Crucially, the impact of these private entities on the audibility and visibility of speech (through actions like deplatforming, demotion, or automated filtering) is now demonstrably greater and more immediate than the effect of any single national law.

This transition represents a shift from a visible, State-centric threat to an invisible, decentralized threat emanating from the private sector, demanding a fundamental re-evaluation of where the true power of censorship lies.

2.1. The Rise of Private Censors: Deplatforming and Content Moderation

In the European context, the guarantees of Article 10 of the ECHR apply primarily to the relationship between the state and the citizen (State Action Doctrine)⁶. However, private platforms are not traditionally considered "public authorities" and are therefore not directly obliged to comply with fundamental rights standards.

This discrepancy creates two major problems: in the first place, deplatforming as a private enforcement, which means that the removal of a user or account from a major platform amounts to an exclusion from the marketplace of public discourse. Platforms make decisions quickly, without the transparency or procedural safeguards of a court. Even if there is an internal appeal, it does not offer the same protections against discretionary censorship as a judicial system⁷.

⁶ The "State Action" gap highlights the theoretical difficulty of applying fundamental rights horizontally (from citizen to private entity). This legal void is the primary driver for new regulatory efforts, such as the EU's Digital Services Act (DSA), which attempts to impose public-law obligations, including due diligence and redress mechanisms, on large platforms (Very Large Online Platforms - VLOPs) to protect user rights against private governance. See European Parliament, *The fight against disinformation and the right to freedom of expression* (Policy Department for Citizens' Rights and Constitutional Affairs, 2021)

⁷ The process of deplatforming is widely criticized for the lack of procedural due process. This includes enforcing vague Terms of Service, failing to provide clear notice or rationale for restrictions, and lacking impartial review. Legal scholars advocate for the imposition of non-discrimination and due process standards—principles historically applied to common carriers: to limit the platforms' unchecked discretion in content moderation decisions. See NUNZIATO, Dawn C., *Protecting Free Speech and Due Process Values on Dominant Social Media Platforms*, George Washington University Law School, 2022.

Likewise, there is also shadow banning⁸ and algorithmic control, which supposes that the threat is no longer just visible (deleting a post), but also invisible. Through the mechanism of shadow-banning or algorithmic adjustment of visibility (ranking), platforms can drastically reduce the audience of a speech or a voice, without an explicit decision. This opaque control represents an attack on the right to visibility, which is essential for the effective exercise of free expression in digital society.

This private power of censorship, wielded by digital gatekeepers, goes beyond mere annoyance, fundamentally altering the nature of online expression. When a single deplatforming decision can instantly disconnect a user from their audience, livelihood, and established social network, the platform ceases to be a mere “website” and functions as an essential public utility. The immediate and disproportionate nature of the sanction (often a permanent ban with no effective recourse) induces a profound “chilling” effect not just on the individual targeted, but on the entire user base. Users, fearing the opaque and heavy hand of the algorithm or content moderator, resort to self-censorship, thus voluntarily narrowing the spectrum of public debate long before a platform policy is invoked⁹.

Platform’s Terms of Use and Community Guidelines¹⁰ thus function as a form of private global law (*Lex Digitalis*). Unlike state-enacted law, these rules are drafted unilaterally, frequently amended, and applied to a global audience of billions of people, often without regard to cultural or legal specificities (e.g., specific national laws on defamation or hate speech). This system of private governance creates an inherent tension between the platform’s self-interest (to maintain an advertiser-friendly environment that is compliant with commercial interests) and the fundamental interest of democracy (to foster a robust, diverse, and even controversial

⁸ Shadow-banning (or “visibility reduction”) is a critical concern because it operates *clandestinely*, limiting the reach of a user’s content without their knowledge, which is conceptually different from outright content removal. Research demonstrates that this opaque algorithmic control can be used to subtly shift public opinion or increase/decrease polarization, even while appearing neutral to external observers. This undercuts the effectiveness of free speech by denying lawful discourse the opportunity to be meaningfully received. See Tauhid Zaman and Yen-Shao Chen, *How Shadow Banning Can Silently Shift Opinion Online*, Yale Insights (2024).

⁹ The classification of dominant platforms as essential public utilities is a core legal debate, supported by their function as “all-powerful gatekeepers.” The lack of due process in their *deplatforming* decisions (often immediate and final) generates a profound chilling effect and encourages self censorship among users, out of fear of social or professional exclusion, thus undermining the principle of a pluralistic marketplace of ideas. See Joseph A. Custer, *Cancel Culture and Censorship Effects Before and After the Introduction of Social Media*, (2024).

¹⁰ The terms of service of platforms are frequently conceptualized as a form of private global law (*Lex Digitalis*), functioning as an alternative legal order that often trumps national jurisdictions and constitutional norms. This system is driven primarily by the platforms’ commercial interests (data exploitation and profit maximization) rather than by the protection of fundamental rights, creating a fundamental tension between the profit-seeking private power and the democratic necessity for an unconstrained public sphere. See Michelle Akim, *Human Rights in the (Social) Media Contract: Digital Constitutionalism’s Response to Government Censorship Online* (McGill University, 2023)

marketplace of ideas). The result is a system in which access to the modern public sphere depends not on compliance with constitutional standards or human rights, but on adherence to the arbitrary and commercially motivated dictates of a few powerful corporations.

2.2. The AI-Driven War on Misinformation and the Chilling Effect

The fight against disinformation and fake news is justified by the legitimate need to protect democratic integrity and public order. However, the tool used to enforce this fight (AI – Artificial Intelligence), which creates a new set of legal challenges by shifting the power of judgment from human oversight to opaque algorithms, amplifying the core issues of bias and due process.

On the one hand, it can be said that AI effectively arbitrages the truth, because AI and Machine Learning (ML) algorithms are increasingly being used to identify, label, and remove content deemed false or harmful at a massive scale. The fundamental problem is that the definition of "truth" or "harm" becomes subjective and embedded in datasets and programming logic of private companies. These systems are often trained on large volumes of labelled human data, which can inherit or even amplify existing prejudices in society. An algorithm trained incorrectly or with unrepresentative data can lead to the suppression of legitimate speech, satire, minority opinions, or inoffensive terms. For example, hate speech detection models have been shown to often exhibit racial and linguistic biases, classifying common phrases or dialects (such as African American Vernacular English - AAVE) as "toxic" at significantly higher rates than other types of language, leading to over-policing and disproportionate deletion of content from marginalized communities¹¹. An error or bias in the algorithm therefore can lead to the suppression of legitimate discourse, satire, or minority opinions, unfairly affecting the groups that the system should, in theory, protect.

On the other hand, the aggressive, rapid, and opaque use of AI to moderate content and penalize disinformation creates an amplified chilling effect. Unlike the censorship applied by a court or a human moderator, the algorithmic decision lacks transparency and speed in providing an explanation or an effective appeal mechanism. When platforms automatically delete content

¹¹ The function of AI as an arbiter of truth is deeply problematic due to the phenomenon of algorithmic bias. Studies show that AI models trained to detect hate speech and toxicity frequently misclassify content based on linguistic and racial markers, disproportionately flagging dialectal speech (e.g., AAVE) or identity terms as toxic. This creates a risk of over-moderation and systematic discrimination against minority groups. Thomas R. P. and W. R., *Racial Bias in Hate Speech and Abusive Language Detection Datasets*, Proceedings of the ACM on Human-Computer Interaction, Vol. 3, No. CSCW (2019) The lack of transparency in the datasets and the inherent difficulty of AI in discerning context, satire, or irony exacerbate this problem.

or reduce the visibility of controversial discourse (shadow-banning), users do not know exactly why they were penalized or how to modify their future expression to conform to the platform's rules. This uncertainty leads to widespread self-censorship, wherein users proactively avoid sensitive topics, legitimate political debate, or criticism for fear of being silently excluded from the system. This invisible and disproportionate mechanism undermines the fundamental function of Article 10 ECHR, which is to allow for a vigorous and unfettered exchange of ideas, by reducing public discourse to an inoffensive common denominator determined by the technological and commercial imperative of the platform.¹²

3. RECONCEPTUALIZING FREE EXPRESSION FOR THE DIGITAL AGE

If threats to free expression have evolved from state oppression to control exercised by private entities, it is not merely desirable but imperative that legal protection instruments adapt. Historically, the fundamental guarantee of free expression, encapsulated in charters such as the European Convention on Human Rights (ECHR) and various national constitutions, was meticulously crafted as a vertical shield for the citizen against the overreaching arm of the State.

This traditional legal dogma, known as the State Action doctrine (or the requirement for immutability in European law), asserts that fundamental rights apply exclusively in the State-Citizen relationship. In this classical framework, private corporations, regardless of their size or influence, are not considered bound by these constitutional restraints. However, the rise of monopolistic online platforms has created an unbridgeable Constitutional Lag where the very infrastructure of public discourse is now governed by private terms of service, not by democratic law.

The digital space is no longer a peripheral territory, but the principal arena for political, social, and economic life. Large technology corporations, wielding the power to silence, amplify, or invisibilize content on a global scale, act, de facto, as "administrators of public space"

¹² The digital chilling effect is intensified by the opaqueness and scale of automated moderation, which leaves users uncertain about the limits of permissible speech. The European Court of Human Rights jurisprudence emphasizes that the right to freedom of expression is crucial for a democratic society, requiring that any interference be prescribed by law and proportionate. The Council of Europe has highlighted that automated content moderation processes must be subject to human oversight and rigorous evaluation to prevent them from unduly restricting users' fundamental rights. See Council of Europe, Parliamentary Assembly, *Regulating content moderation on social media to safeguard freedom of expression* (Doc. 15729, December 2024); see also T. McGonagle, *The concept of chilling effect: Its untapped potential to better protect democracy, the rule of law, and fundamental rights in the EU* (Open Society Foundations, 2021).

or "digital sovereigns". They exercise a form of unconstrained governance over billions of users, a power that rivals, and often exceeds, that of many nation-states.

Consequently, the traditional legal mechanism can no longer support effective protection; a profound rethink is needed to transition from a purely vertical application of rights to a model of horizontal applicability, thereby imposing fundamental rights standards (particularly transparency, due process, and non-discrimination) also on these dominant private entities. Failure to do so renders the constitutional promise of free speech illusory in the twenty-first century.

3.1. The Right to Visibility: Algorithmic Opaque Governance

A crucial aspect of this reimagining is to extend protection beyond the simple right to publish (the right to upload content) to include the right to visibility (the right for that content to have a reasonable chance of being seen).

In the contemporary, saturated digital marketplace, the mere physical right to speak means nothing if your speech is instantly silenced (or rendered invisible) by opaque algorithmic filters. These algorithms, particularly those governing content ranking and recommendation, act as the new gatekeepers of public discourse, deciding what content is amplified and what is marginalized—a phenomenon often described as the Algorithmic Cage. Without this visibility, the guaranteed right to speak fundamentally devolves into a "right to whisper into the void"¹³.

Recognizing this Right to Visibility (*droit à la visibilité*) would necessitate a fundamental shift in platform accountability. It would require platforms to justify the design and deployment of their recommendation algorithms, demonstrating that they are not inherently discriminating against lawful speech based on internal biases (e.g., favouring sensationalist content over reasoned debate, or penalizing specific political viewpoints). Furthermore, the Right to Visibility implies a form of algorithmic due process, demanding that users be informed when their content's reach is significantly reduced by demotion or shadow-banning and be given an effective, transparent mechanism to challenge that automated decision.

3.2. The Judicial Response: Tackling the State Action Doctrine

¹³ The concept of the Right to Visibility (or Right to be Heard) is a conceptual necessity to adapt Article 10 ECHR to the digital environment. In a platform-dominated public sphere, the right to post is meaningless if content is systematically demoted or hidden by non-transparent algorithmic prioritization systems (ranking). This right recognizes that effective freedom of expression requires a reasonable chance of reception by the intended audience, making algorithmic processes a central concern for human rights law. See SINE, N., *Controversial Encounters in the Age of Algorithms: How Digital Technologies Are Stifling Public Debate and What to Do About It*, 2025.

Faced with this new reality, the international and national judicial system are struggling to adapt the classic mechanisms.

The European Court of Human Rights (ECHR), constrained by its mandate to adjudicate state actions, has in some cases attempted to extend the state's positive obligation to ensure a legislative framework that protects users from abuses by private platforms. The Court holds that states must take reasonable steps to protect the right to expression from interference by private entities where the platform functions as an essential means of political participation or social discourse. This approach attempts to bridge the gap by shifting the burden of regulation onto the state, thus indirectly holding private platforms accountable for human rights standards¹⁴.

Yet, the European Union provides the most decisive and far-reaching response at the global level through the adoption of the Digital Services Act (DSA) (Regulation (EU) 2022/2065). This landmark legislation represents a paradigm shift, compelling large online platforms to function within a framework akin to public accountability. The DSA introduces two pivotal mechanisms.

First, it enforces increased transparency, often referred to as algorithmic accountability. Very large online platforms (VLOPs) must disclose to both users and authorities detailed information about how their content moderation algorithms and recommendation systems operate. This provision directly counters the so-called “Algorithmic Cage”, allowing independent audits and external scrutiny to uncover systemic biases and potential risks.

Second, the DSA establishes objection and redress mechanisms grounded in principles of due process. Platforms are required to implement faster, fairer, and genuinely independent internal appeal procedures, enabling users to contest decisions related to content removal or de-platforming. Significantly, the DSA also ensures access to out-of-court dispute resolution bodies, offering users a pathway to seek remedies beyond the control of the platform itself. This is a critical move toward holding private entities accountable to due process standards.¹⁵

¹⁴ The ECHR doctrine on positive obligations requires states to take legislative and administrative steps to ensure that fundamental rights, including freedom of expression, are respected even in the sphere of private relations. While the Court cannot directly regulate private platforms, it holds the State responsible for failing to adopt a regulatory framework that provides adequate protection and remedies for users against arbitrary decisions by powerful platform operators. See ECHR, *Case law on content moderation and freedom of expression* (Special Collection, November 2024).

¹⁵ The Digital Services Act (DSA) is a landmark legislative attempt to impose public law standards onto private digital infrastructure. By mandating increased algorithmic transparency and establishing mandatory, accessible, and fair complaint-handling systems (Art. 17-20), the DSA effectively treats Very Large Online Platforms (VLOPs) as entities with quasi-public responsibilities. This regulatory shift aims to mitigate the Lex Digitalis (private law) by subordinating commercial content moderation practices to a regime that prioritizes transparency,

Through the DSA, the EU makes the first substantial, globally meaningful attempt to compel major platforms to adhere to fundamental rights principles, bridging the gap between constitutional guarantees and the realities of private digital governance. These robust tools (transparency, oversight, and effective mechanisms for redress) constitute the essential first step in dismantling the walls of the Algorithmic Cage.

CONCLUSION AND PROSPECTS

This essay has proved that the nature of threats to Freedom of Expression has undergone a fundamental metamorphosis. If in the recent past, symbolically illustrated by the years of the Iron Curtain, constitutional guarantees were violated by the brutal political will of the state, today, the gap between the normative guarantee and the legal reality is generated by much more subtle and less transparent forces: algorithms and private Digital Governance.

The central thesis was supported by three major arguments. First, we showed how the power of censorship has shifted from public institutions to large technology corporations, with the phenomenon of de-platforming becoming the modern equivalent of exclusion from public space. Second, we analysed the risk that Artificial Intelligence (AI), used in the fight against disinformation, could generate a strong chilling effect on legitimate and minority discourse. Finally, it became clear that the classical legal doctrine of State Action is inadequate to hold these private entities accountable, thus leaving fundamental law vulnerable.

Reimagining the protection of Fundamental Rights is no longer a theoretical option, but a pragmatic necessity. Legislative solutions, such as the European Union's Digital Services Act (DSA), represent a bold and essential first step to impose algorithmic transparency and fair redress mechanisms against private censors. However, an evolution of jurisprudence is also needed that recognizes, implicitly or explicitly, the Right to Visibility: the right for lawful discourse to have a reasonable chance of being heard in the digital environment.

In conclusion, the fight for freedom of expression is not over; it has simply moved to a new front. The Algorithmic Cage is less visible than the physical walls of the past, but its threat to the diversity of ideas and, implicitly, to democracy, is just as profound. The future of constitutional guarantees depends on the ability of the legal system to adapt and extend the protection of rights beyond the physical borders of the state, into the virtual territory of private power. Only in this way can we ensure that the right to speak freely does not ultimately turn into a simple right to whisper into the void.

accountability, and the fundamental rights of users. *See* Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services (Digital Services Act), Art. 17-20

BIBLIOGRAPHY

I. Doctrine

1. Akim, Michelle, *Human Rights in the (Social) Media Contract: Digital Constitutionalism's Response to Government Censorship Online* (Master's Thesis, McGill University, 2023).
2. Corobca, Liliana, *Controlul Cărtii. Cenzura Literaturii în Regimul Comunist din România*, Editura Cartea Românească, 2014.
3. Deletant, Dennis, *Ceausescu and the Securitate: Coercion and Dissent in Romania, 1965-1989*, C. Hurst & Co., 1995.
4. Vasile, Cristian, *Literatura și artele în România comunistă, 1948-1953*, Editura Humanitas, 2010.

II. Specialized articles and studies

1. Council of Europe, Parliamentary Assembly, *Regulating content moderation on social media to safeguard freedom of expression* (Doc. 15729, December 2024).
2. Custer, Joseph A., *Cancel Culture and Censorship Effects Before and After the Introduction of Social Media* (2024).
3. European Parliament, *The fight against disinformation and the right to freedom of expression* (Policy Department for Citizens' Rights and Constitutional Affairs, 2021).
4. P., Thomas R., and R., W., *Racial Bias in Hate Speech and Abusive Language Detection Datasets*, Proceedings of the ACM on Human-Computer Interaction, Vol. 3, No. CSCW (2019).
5. SINE, N., *Controversial Encounters in the Age of Algorithms: How Digital Technologies Are Stifling Public Debate and What to Do About It* (2025).

III. Sources in electronic format

1. McGonagle, Tarlach, *The concept of chilling effect: Its untapped potential to better protect democracy, the rule of law, and fundamental rights in the EU*, Open Society Foundations, 2021, available at: <https://www.opensocietyfoundations.org/uploads/...>
2. NUNZIATO, Dawn C., *Protecting Free Speech and Due Process Values on Dominant Social Media Platforms*, George Washington University Law School (2022), available at: https://scholarship.law.gwu.edu/faculty_publications/1589/
3. Zaman, Tauhid, and Chen, Yen-Shao, *How Shadow Banning Can Silently Shift Opinion Online*, Yale Insights (2024), available at: <https://insights.som.yale.edu/insights/...>

IV. Case law

1. European Court of Human Rights, *Case law on content moderation and freedom of expression* (Special Collection, November 2024).

V. Legislation

1. Constituția României (The Constitution of Romania), Art. 30(1).
2. European Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR), 1950, Article 10.

-
3. Law no. 3 of 28 March 1974 on the press in the Socialist Republic of Romania, available at: <https://legislatie.just.ro/Public/DetaliuDocument/338>.
 4. Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act, DSA), Art. 17-20.