

SOCIAL MEDIA AS A HABITAT FOR HYBRID THREATS: A COMPARATIVE ANALYSIS OF COMMUNICATION MODELS FOR PLATFORM REGULATION IN THE EU, THE U.S., AND CHINA

Nadiia Balovsiak

Ph.D. in Pedagogy, Associate Professor at the School of Journalism and Communications,
Ukrainian Catholic University, Ukraine

e-mail: balovsyak@ucu.edu.ua, orcid.org/0000-0002-1810-7397

Summary

Social networks are increasingly viewed not as a neutral channel of communication, but as a structurally vulnerable media environment in which their architectural features – engagement algorithms, cross-border infrastructure, and minimal barriers to publication – facilitate the spread of manipulative content. This communicative vulnerability makes platform regulation not an optional tool of media policy, but a structural necessity. The aim of this article is to conduct a comparative analysis of three models of platform governance – those of the European Union, the United States of America, and the People’s Republic of China – through the lens of media theory, public sphere theory, and the communicative power of platforms. The study reveals that despite fundamental differences in approaches – the EU’s rights-based media policy, the U.S.’s security-oriented logic, and China’s sovereignty-based control – all three models face a common structural deficit: regulatory mechanisms systematically lag behind the pace of transformation in communication practices and technologies of manipulative influence. The article proposes a typology of communicative regulatory models (rights-based, security-based, and sovereignty-based) and argues that none of them eliminates the fundamental asymmetry between the speed of technological change and the speed of institutional response to these changes, but merely distributes the associated communicative risks and costs in different ways. The findings of this study are significant for the development of media policies in countries that, like Ukraine, are simultaneously targets of hybrid information threats and actors in shaping their own architecture of the digital communication space.

Key words: manipulative communication, media environment, platform governance, algorithmic architecture, disinformation, information sovereignty, public sphere, media literacy.

DOI <https://doi.org/10.23856/7604>

1. Introduction

Social networks are increasingly viewed not as a neutral channel of communication, but as a structurally vulnerable media environment – one in which the very architecture of the platform creates conditions for manipulative influence, particularly hybrid information threats. Regulating this environment faces a problem structurally similar to that described by media environment theory: a technological platform is not a neutral channel for transmitting messages – it shapes the content and conditions of communicative interaction. While Marshall McLuhan formulated this principle in the context of traditional media, in the context of algorithmic platforms it takes on a new dimension: the medium is not merely “the message” (McLuhan, 1964); it actively constructs hierarchies of visibility, determines which messages

will be seen by as many people as possible, and thereby becomes an independent actor in public communication.

Social networks are not merely a channel through which hybrid threats spread, but a media environment that structurally enhances their communicative effectiveness. Recommendation algorithms optimized for attention retention, minimal barriers to publication, and the cross-border nature of the infrastructure create conditions under which emotionally charged, polarizing content spreads faster and reaches a wider audience than would be possible under the traditional model of editorial media. This is first and foremost a communicative phenomenon – and it is precisely communicative analysis that can reveal its structural nature, which legal or security approaches merely document but do not explain.

The relevance of this study is determined by several interrelated factors. First, the scale of these platforms' communicative influence: as of the end of 2025, the combined audience of the largest social networks exceeds 5.4 billion monthly active users (*DataReportal / We Are Social, 2025*), making them the largest media platforms in human history. Second, the documented role of social media in electoral and geopolitical crises over the past decade – from interference in the 2016 U.S. election to the systematic use of these platforms in a full-scale information war against Ukraine since 2022. This convincingly demonstrates that their vulnerability to manipulative use directly threatens public discourse and democratic institutions. Third, the three leading global jurisdictions have adopted fundamentally different responses to this vulnerability, and a comparative analysis of their approaches is crucial for the development of media policy strategies by states at various stages of digital regulation, particularly Ukraine, which is simultaneously the target of intense hybrid information threats and an actor in shaping its own architecture of the digital communicative space.

The purpose of this article is to conduct a comparative analysis of three models of platform governance (the EU, the U.S., and China) through the lens of their approach to the communicative architecture of platforms: regulation of algorithms, content, ownership structure of infrastructure, or access to the public information space. Research question: Is there a common regulatory vulnerability across all three models, despite their fundamental differences? The methodological basis of the article is a comparative analysis of three case studies using a unified matrix: the object of regulation, the mechanism for implementing regulatory decisions, the speed of response to new threats, and vulnerability to external pressure.

2. The Algorithm as Editor

In this study, the terms “social networks” and “algorithmic platforms” are used synonymously – both refer to digital communication environments whose architecture is based on algorithmic content ranking. The theory of media ecology, developed in the works of McLuhan and Postman, provides the conceptual framework necessary to understand why social networks constitute a qualitatively distinct type of communicative space. Media are not neutral containers for content – they actively shape the cognitive and social practices that take place through them. Postman expanded on this idea, arguing that every medium carries its own “epistemology” – a way of knowing and representing reality (*Postman, 1985*). Algorithmic platforms, in this sense, embody an “epistemology of engagement”: the criterion of relevance is not truth or social significance, but the ability to hold attention for as long as possible and generate an emotional response.

Unlike traditional media, where an editorial team with specific journalistic standards stands between the author and the audience, an algorithmic platform replaces the editor with an

automated ranking system. In such an environment, producers of manipulative content gain a structural advantage: emotionally charged, polarizing posts align better with algorithmic logic than measured, balanced content. Empirical research confirms this pattern – moral and emotional content spreads on social media significantly faster than neutral posts (*Brady et al., 2017*) means that the platform is not a neutral intermediary in public communication – its architecture systematically favors precisely the type of posts that elicit the strongest emotional response.

The concept of “filter bubbles” (*Pariser, 2011*) describes yet another communicative effect of algorithmic architecture: the personalization of the information flow leads to different groups of users receiving fundamentally different versions of public reality. This effect is directly related to the problem of hybrid threats: an actor seeking to polarize society not only disseminates destructive content but also relies on the platform’s algorithmic logic itself, which segregates audiences and minimizes their exposure to alternative viewpoints. Thus, the platform becomes not merely a channel but an active participant in polarizing communication, even without any intention on the part of its owners.

Habermas’s concept of the public sphere as a space for rational communicative interaction (*Habermas, 1989*) undergoes a fundamental transformation in the context of algorithmic platforms. In classical theory, the public sphere functions according to the logic of rational debate: the legitimacy of public discourse is determined by the quality of the argument, not by the status of the person expressing it. An algorithmic platform replaces this principle with the principle of maximum engagement: visibility in the public sphere is determined not by the quality of the argument, but by its ability to generate reactions. This transformation has direct implications for regulation: if the problem lies not in specific “bad” content but in the very architecture of the communicative environment, then the solution requires intervention in that architecture, not merely in the content.

The factor of cross-border nature is added to the algorithmic logic. Social platforms operate simultaneously across dozens of jurisdictions, while states’ regulatory powers are limited to their own territory. An actor carrying out hybrid information influence through a platform while physically located outside the target state remains virtually untouchable – and the very communication environment through which this influence is exerted continues to function without restrictions.

Let’s examine each of the three models in more detail, starting with the European one.

3. The European Union: Regulating Systemic Risks

The Digital Services Act (*DSA; European Commission, 2022*) is the most comprehensive attempt among the three models analyzed to regulate not individual pieces of content, but the communicative architecture of platforms as such. The DSA imposes on the largest platforms – so-called “very large online platforms” (VLOPs) with an audience of over 45 million users in the EU – the obligation to conduct regular assessments of systemic risks and take measures to mitigate them (Articles 34–35 of the DSA). Among the categories of systemic risks identified by the DSA, the following are singled out: foreign information manipulation and interference (FIMI), algorithmic amplification of harmful content, and threats to media literacy and access to diverse information.

From the perspective of media environment theory, the DSA represents the regulator’s first major attempt to intervene directly in a platform’s “epistemology” – that is, in the mechanisms that determine what information becomes visible and to whom. Article 38 of the DSA requires platforms to provide users with alternative recommendation systems that are not based

on user profiling. This is a direct intervention in the architecture of the personalized communication environment, which is a structural prerequisite for “filter bubbles” (*Pariser, 2011*). The requirement for transparency in recommendation algorithms (Article 27) opens the door for researchers to verify whether an algorithm amplifies manipulative content – thereby transforming regulation from a purely administrative mechanism into a tool for communicative audit.

Enforcement of the DSA is gaining real momentum. In December 2025, the European Commission imposed the first major fine for a DSA violation – 120 million euros on the X platform for three breaches of transparency obligations: the misleading design of the paid account verification system (“blue checkmark”) (Article 25 of the DSA), insufficient transparency in the advertising repository (Article 39 of the DSA), and refusal to grant researchers access to public data (Article 40 of the DSA) (*European Commission, 2025*). This move marks a transition from the stage of formal implementation of the law to the stage of actual communicative oversight of platforms.

At the same time, this model reveals a significant structural limitation: its effectiveness critically depends on the regulator’s willingness and ability to resist geopolitical pressure from the countries where the largest platforms are registered. The implementation of the DSA was accompanied by an open conflict with the United States: the U.S. administration imposed visa restrictions on a number of European officials involved in the development and enforcement of the law, and tariff policy was used as a tool to exert pressure regarding concessions in digital regulation (*The Irish Times / Financial Times, 2026*). This demonstrates a fundamental vulnerability in the rights-based regulatory model: even a legally robust set of tools proved dependent on the balance of geopolitical power between the regulator and the country where the largest platforms are registered.

A separate limitation is temporal asymmetry – the gap between the speed of technological change and the pace of institutional response. The DSA took several years to be drafted and adopted before it came into effect. During this time, communication technologies for manipulative influence – in particular, generative artificial intelligence and synthetic media – underwent several generations of development, and the law required adaptation even before its actual enforcement began.

4. United States: Security Through Infrastructure Control

The American model of platform regulation reflects a specific constitutional and cultural attitude toward freedom of communication: The First Amendment traditionally protects content from government interference without imposing obligations regarding the quality of the communication environment. The result is a legal framework that exempts platforms from liability for content posted by third parties (Section 230 of the Communications Decency Act – a provision that has effectively precluded systematic regulation of algorithmic architecture, similar to the DSA). In this culture, the regulation of public communication is traditionally viewed as a threat to free speech rather than as a means of protecting it. Section 230 of the Communications Decency Act (1996) became the legal foundation that enabled the development of social media in their current form: platforms are not liable for content posted by their users, and can therefore scale up without editorial oversight (*Communications Decency Act, 1996*). It is precisely this provision that has allowed platforms to act simultaneously as passive intermediaries – which relieves them of legal liability – and as active architects of the communicative environment through algorithmic ranking. This ambiguity has become central to discussions about reforming U.S. platform regulation.

In response to growing concerns about foreign manipulative influence via social platforms, the U.S. model has adopted an unconventional solution: rather than regulating the communication architecture, it regulates ownership of the communication infrastructure. The Protecting Americans from Foreign Adversary Controlled Applications Act (*PAFACA*, 2024) provides a procedure under which the president may require a platform controlled by an entity from a “hostile” jurisdiction to either divest (forcibly sell a stake in ownership) or prohibit its availability in U.S. app stores. The Supreme Court upheld the constitutionality of this approach in *TikTok Inc. v. Garland* (2025), ruling that the law is content-neutral – since it targets a foreign adversary’s control over a company rather than specific content or viewpoints – and applying the standard of intermediate constitutional scrutiny to it.

From the perspective of communication studies, this approach raises an important theoretical question: Is the owner of a communication infrastructure the determining factor in communicative influence? The American logic suggests that it is: if the Chinese company ByteDance controls TikTok’s algorithm, then it potentially controls which content is prioritized in American users’ feeds. This logic is entirely consistent with media environment theory: the platform’s architecture shapes the communicative environment – yet the proposed solution to the problem is to replace the owner, rather than to change the architecture itself.

From a communicative perspective, the structural limitation of this model lies in its inherent selectivity. The algorithmic vulnerability of the communicative environment – the ability of any platform, regardless of its origin, to serve as a medium for manipulative influence through the mechanism of engagement – remains outside the scope of regulatory attention. Even the complete displacement of foreign-owned platforms does not eliminate the structural problem: an American company that optimizes its recommendation algorithm to maximize engagement reproduces the same communicative effects as a foreign one. The criterion for regulatory scrutiny is not the architecture of the communicative environment, but the nationality of its owner.

At the same time, the American model demonstrates its own version of temporal asymmetry – but of a different type than in the EU. The TikTok case went through several court instances in less than a year and concluded with a Supreme Court ruling in January 2025. However, every month of legal proceedings is time during which a platform with a potentially problematic owner continued to shape the communication environment for millions of users. Legal proceedings, as a tool for regulating communication risks, are structurally slower than the risks themselves.

5. China: Sovereign Control of Communication

The Chinese model is based on the principle of cybersovereignty – complete state control over the digital communication space even before content becomes publicly available. From a theoretical standpoint, this is the most radical response to the problem of algorithmic vulnerability: if a platform constructs a communication environment, then the state, which controls all platforms, also controls the entire environment. The 2022-2025 regulatory framework – Regulations on the Regulation of Recommendation Algorithms (2022), Regulations on Deep Synthesis (2023), Temporary Measures for the Management of Generative Artificial Intelligence Services (2023), and Measures for the Labeling of AI Content (effective September 2025) – form a multilevel system of control over the communicative architecture of platforms (*Arcesati & Triolo*, 2023).

A telling example is the approach to the issue of synthetic content. The measures for labeling AI-generated content (*Cyberspace Administration of China*, 2025) mandate visible

(to the user) and metadata-embedded labeling for all AI content – text, images, audio, video, and virtual scenes. Unlike similar requirements in the EU’s AI Act (Article 50) (*European Parliament, 2024*), where the labeling obligation falls primarily on the developers of generative models, the Chinese model extends responsibility to the entire communication chain: the model developer, the distribution platform, and the user themselves, who is required to declare the use of generative technologies upon publication. This is an example of systems thinking in the regulation of communication processes: responsibility is distributed among all actors involved in the production and dissemination of media content.

The Provisions on the Regulation of Recommendation Algorithms (2022) (*Cyberspace Administration of China, 2022*) are even more telling: it establishes an obligation for platforms not to use algorithmic recommendations to disseminate content that “violates public morality,” “threatens public order,” or “undermines national unity.” These formulations are extremely broad – and it is precisely this breadth that is a key feature of the Chinese model’s communicative architecture: the regulator retains maximum flexibility in defining the boundaries of permissible public discourse.

A structural limitation of the Chinese model is that it does not distinguish between protecting public discourse from external manipulation and state manipulation of that very discourse. Preemptive control over the communication environment effectively blocks external influences – but that same infrastructure can be used to manipulate public opinion from within, and it is impossible to independently verify this. The decision to allow content into the public sphere is made by the very same body that has the authority to exert such influence. This does not solve the problem of hybrid communication threats; rather, it represents the state’s monopolization of the issue.

At the same time – and this is a significant analytical point – the Chinese model demonstrates the fastest rate of adaptation to new communication technologies among the three systems studied. Less than two years passed between the emergence of large language models and the introduction of mandatory labeling of AI-generated content in China. This speed is explained not only by the authoritarian nature of the system but also by its specific institutional architecture: the Cyberspace Administration of China (CAC) is the sole supervisory body with broad regulatory and law enforcement powers, which eliminates the problem of interagency coordination characteristic of the EU and the U.S.

6. Comparative Analysis: Common Structural Patterns

Comparing the three models using a single analytical matrix reveals fundamental differences in what exactly becomes the focus of regulatory attention. The European model regulates a platform’s communication architecture – algorithms, systemic risks to public discourse, and the transparency of recommendation systems – regardless of the platform’s geographic origin. The American model regulates the owner of the communication infrastructure, but not the architecture itself. The Chinese model regulates content access to the public communication space in a preventive and comprehensive manner, even before the content is disseminated.

Despite these differences, all three models exhibit a common structural shortcoming that can be termed a “regulatory temporal gap”: none of them resolves the fundamental mismatch between the pace of transformation in communication technologies and the pace of the institutional response. This pattern is structural, not accidental: it recurs regardless of the political nature of the system or the specific object of regulation. Generative artificial intelligence, which

in just a few years has fundamentally changed the possibilities for creating synthetic content, is merely the most recent example of this pattern – a similar situation was observed with the emergence of bots, microtargeting, and deepfakes.

It is also important to note that, in the communicative dimension, models are not airtight and interact with one another in ways that can themselves become vulnerabilities. Discrepancies between different approaches to regulating the platform environment are actively exploited by actors interested in weakening regulatory protections for public discourse. The open conflict between the U.S. and the EU over the DSA is illustrative here: geopolitical pressure is being used not only for economic reasons but also to limit the EU's ability to regulate the communicative architecture of platforms, which is a strategic asset of American technology companies.

From the perspective of public sphere theories, all three models reflect different answers to the same question: who is the “editor” of public discourse in the digital age? The liberal tradition (EU/U.S.) seeks to distance the state as much as possible from this role and place responsibility on platforms through transparency and accountability. The Chinese model explicitly designates the “state” and makes this choice a foundational principle of digital media policy. None of these answers is neutral regarding the nature of public discourse as it takes shape within the respective environment.

7. Conclusions

First, the algorithmic architecture of social networks is a structural – rather than a random – factor in the platforms' communicative vulnerability to hybrid threats. Recommendation algorithms optimized for engagement systematically favor content types that align with the logic of manipulative influence, regardless of the intentions of specific actors. The transformation of the public sphere in the context of algorithmic platforms is fundamental: the criterion for visibility in public discourse has shifted from the quality of an argument to the ability to generate an emotional response. This means that effective regulation must focus primarily on the architecture of the communicative environment, rather than merely on individual pieces of content.

Second, all three models examined – the human rights-based (EU), security-based (U.S.), and sovereignty-based (PRC) – despite fundamental differences in the object and mechanism of regulation, exhibit a common structural shortcoming: the lag in the regulatory response relative to the pace of technological transformation in communication practices is a pervasive pattern, rather than a localized shortcoming of any specific model. None of the models eliminates the fundamental asymmetry between the speed of change and the speed of the institutional response – they merely distribute the associated communicative risks and costs in different ways.

Third, the article outlines two directions for further research. The first is an analysis of how states that, like Ukraine, are simultaneously the target of intense hybrid information threats and the agents of developing their own digital communicative space, shape media policy – specifically, how they shape their own media policy amid active information confrontation. The second is an examination of the role of media literacy and critical media consumption as factors in the public sphere's resilience to manipulative communicative influences: regulatory mechanisms focused exclusively on platforms, without corresponding investments in communicative education for audiences, address only part of the structural problem.

References

1. Arcesati, R., & Triolo, P. (2023). *China's AI regulations and how they get made*. Carnegie Endowment for International Peace. <https://carnegieendowment.org/research/2023/07/chinas-ai-regulations-and-how-they-get-made>
2. Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences*, 114(28), 7313–7318. <https://doi.org/10.1073/pnas.1618923114>
3. Cyberspace Administration of China. (2022). *Provisions on the Administration of Algorithm-Generated Recommendations for Internet Information Services*. <https://www.chinalaw-translate.com/en/algorithms/>
4. Cyberspace Administration of China. (2025). *Measures for the Labeling of AI-Generated and Synthetic Content*. <https://www.chinalawtranslate.com/en/ai-labeling/>
5. DataReportal / We Are Social. (2025). *Digital 2025: Global overview report*. <https://datareportal.com/reports/digital-2025-global-overview-report>
6. European Commission. (2022). *Regulation (EU) 2022/2065 of the European Parliament and of the Council on a Single Market for Digital Services (Digital Services Act)*. Official Journal of the European Union. <https://eur-lex.europa.eu/eli/reg/2022/2065/oj>
7. European Commission. (2025, December 5). *Commission fines X €120 million under the Digital Services Act* [Press release]. https://ec.europa.eu/commission/presscorner/detail/en/ip_25_2934
8. European Parliament. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council (Artificial Intelligence Act)*. Official Journal of the European Union. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689>
9. Habermas, J. (1989). *The Structural Transformation of the Public Sphere*. MIT Press.
10. McLuhan, M. (1964). *Understanding Media: The Extensions of Man*. McGraw-Hill.
11. Pariser, E. (2011). *The Filter Bubble: What the Internet Is Hiding from You*. Penguin Press.
12. Postman, N. (1985). *Amusing Ourselves to Death: Public Discourse in the Age of Show Business*. Viking.
13. *Protecting Americans from Foreign Adversary Controlled Applications Act*, Pub. L. No. 118-50, Div. H (2024).
14. *Communications Decency Act*, 47 U.S.C. § 230 (1996). <https://www.law.cornell.edu/uscode/text/47/230>
15. *The Irish Times / Financial Times*. (2026, January 5). *EU readies tougher tech enforcement in 2026 as Trump warns of retaliation*. *The Irish Times*. <https://www.irishtimes.com/technology/big-tech/2026/01/05/eu-readies-tougher-tech-enforcement-in-2026-as-trump-warns-of-retaliation/>
16. *TikTok Inc. v. Garland*, 604 U.S. ____ (2025).