DOI 10.21564/2225-6555.2026.29.360305

Algorithmic Platform Moderation and Freedom of Expression in Ukraine

Alona D. Harahata*
Yaroslav Mudryi National Law University
Kharkiv, Ukraine
**e-mail: a.d.garagata@nlu.edu.ua*

Abstract
The article examines the transformation of freedom of expression under private algorithmic governance of the public digital sphere. The relevance of the topic is that global online platforms increasingly act as de facto intermediaries for political communication, news exchange, documentation of war crimes, and civic mobilisation, while Ukrainian legislation lacks a comprehensive model for their procedural accountability. The purpose of the article is to identify the legal risks of algorithmic moderation for freedom of expression and to substantiate a regulatory model suitable for Ukraine in the context of war, information aggression, and European integration. The methodology is based on doctrinal, comparative legal, functional, and systemic methods, as well as on the analysis of legal and policy materials concerning the European Union, the United States, the United Kingdom, Taiwan, and Ukraine. The results show that the European Union model provides the most developed procedural safeguards; the American model preserves excessively broad private platform discretion; the British approach emphasises risk-oriented service safety; and the Taiwanese model demonstrates the effectiveness of social resilience and non-punitive coordination. The article substantiates the need for a hybrid model for Ukraine that combines reasoned moderation decisions, internal and independent appeals, systemic risk assessment, an institutional channel of communication with platforms, and rapid response to disinformation, without granting the state censorship powers. Further research should focus on the development of a special law on digital services and platform governance.

Keywords: digital sovereignty; content governance; disinformation; procedural safeguards; systemic risks.

Алгоритмічна модерація платформ і свобода вираження в Україні

Альона Дмитрівна Гарагата*
Національний юридичний університет імені Ярослава Мудрого
Харків, Україна
**e-mail: a.d.garagata@nlu.edu.ua*

Анотація
У статті досліджено трансформацію свободи вираження в умовах приватного алгоритмічного управління публічним цифровим простором. Актуальність теми зумовлена тим, що глобальні онлайн-платформи дедалі частіше виконують функції фактичних посередників політичної комунікації, новинного обміну, документування воєнних злочинів та громадської мобілізації, тоді як українське законодавство не містить цілісної моделі їхньої процедурної відповідальності. Метою дослідження є визначення правових ризиків алгоритмічної модерації для свободи вираження та обґрунтування моделі регулювання, придатної для України в умовах війни, інформаційної агресії та європейської інтеграції. Методологічну основу становлять доктринальні, порівняльно-правові, функціональні й системні методи, а також аналіз правозастосовних і політико-правових матеріалів щодо Європейського Союзу, Сполучених Штатів, Сполученого Королівства, Тайваню та України. Результати дослідження показують, що модель Європейського Союзу забезпечує найрозвиненіші процедурні гарантії; американська модель зберігає надмірно широку приватну дискрецію платформ; британський підхід акцентує ризик-орієнтовану безпеку сервісів; а тайванська модель демонструє ефективність суспільної стійкості й некаральної координації. Обґрунтовано доцільність гібридної моделі для України, яка поєднує мотивування рішень про модерацію, внутрішнє та незалежне оскарження, оцінку системних ризиків, інституційний канал взаємодії з платформами та швидке реагування на дезінформацію без надання державі цензурних повноважень. Подальші дослідження мають бути спрямовані на розроблення спеціального закону про цифрові послуги та платформне управління.

Ключові слова: *цифровий суверенітет; управління контентом; дезінформація; процедурні гарантії; системні ризики.*

Introduction

Freedom of expression remains one of the core guarantees of constitutional democracy. In Ukrainian constitutional law, it is protected by Art. 34 of the Constitution of Ukraine, which guarantees the right to freedom of thought and speech and the free expression of views and beliefs [1]. In the European human rights system, the same guarantee is protected

by Art. 10 of the Convention for the Protection of Human Rights and Fundamental Freedoms [2]. The case law of the European Court of Human Rights has consistently treated freedom of expression as one of the essential foundations of a democratic society, while also recognising that the internet has become a key infrastructure for public debate. This is visible, inter alia, in Delfi AS v. Estonia and Cengiz and Others v. Turkey, where the Court addressed the responsibility of online intermediaries and the importance of access to online platforms for political and social expression [3; 4].

The classical legal model of freedom of expression was built mainly in a vertical paradigm: an individual is protected from excessive interference by the state. However, the contemporary digital environment has changed the structure of communicative power. The practical boundary between visible and invisible speech, permitted and demonetised content, legitimate documentation, and prohibited graphic material is now frequently drawn not by courts or public authorities, but by private platforms through terms of service, community standards, and algorithmic moderation. Recent legal scholarship, therefore, increasingly treats algorithmic moderation not as a purely technical operation, but as a form of governance capable of affecting democratic discourse and fundamental rights [5]. Human rights organisations have also stressed that content moderation must be assessed by reference to legality, legitimacy, necessity, proportionality, transparency, and access to remedy, rather than only by reference to private contractual standards [6].

For Ukraine, this issue is not merely theoretical. Since the beginning of the full-scale Russian invasion, social networks, video platforms, and messengers have become channels for informing civilians, coordinating volunteer assistance, exposing disinformation, documenting war crimes, and maintaining international attention to Russian aggression. At the same time, Ukrainian media, and civil society organisations have reported wrongful restrictions on war-related content, and insufficient sensitivity of global platforms to the Ukrainian context [7]. Freedom House also noted that social media platforms, and search engines removed or restricted war-related content during the invasion, and that several Ukrainian social media pages sharing information about the war or organising support for the army were removed or limited [8].

The relevance of this study is therefore determined by the collision between three regulatory needs. First, Ukraine must preserve constitutional freedom of expression, and avoid transforming platform regulation into state censorship. Secondly, it must develop remedies against arbitrary or contextually incorrect moderation by transnational platforms. Thirdly, it must protect the information environment against hostile manipulation,

disinformation campaigns, and the weaponisation of platform architecture during war. These needs cannot be satisfied by the existing Ukrainian framework alone. The Law of Ukraine On Information and the Law of Ukraine On Media establish important general principles for information relations and media regulation, but they do not create a comprehensive legal status for very large online platforms, nor do they provide a special system of notice, explanation, appeal, risk assessment, and institutional cooperation in cases of algorithmic moderation [9–11].

The purpose of the article is to determine how algorithmic platform moderation affects freedom of expression and to substantiate a hybrid model of platform governance for Ukraine. To achieve this purpose, the article sets the following objectives: to clarify the theoretical shift from state interference to private platform governance; to compare the European Union, United States, United Kingdom, and Taiwanese approaches; to identify the main gaps in Ukrainian law; and to formulate legislative proposals that would combine procedural accountability, systemic risk assessment, institutional cooperation, and non-punitive resilience mechanisms.

Literature Review

The theoretical foundation of the study is formed by the scholarship on digital constitutionalism, platform governance, and algorithmic moderation. Celeste conceptualises digital constitutionalism as a normative response to the concentration of power in the digital environment and argues that constitutional principles must be reconsidered in light of private digital infrastructures [12]. Suzor similarly maintains that platforms exercise rule-making and enforcement functions over users and that the legitimacy of this governance should be assessed through the rule of law, including transparency, predictability, and procedural fairness [13]. These approaches are particularly relevant for the present study because they make it possible to move beyond the narrow contractual understanding of platform-user relations.

Gillespie's work is important because it shows that moderation is not a peripheral or exceptional activity of platforms, but a constitutive function of platform governance [14]. Platforms do not merely host expression; they classify, rank, amplify, suppress, and monetise it. Balkin develops this point from the perspective of freedom of speech theory. He distinguishes older models of speech regulation, centred on state restrictions, from new-school speech regulation, where states and private infrastructure providers interact in complex systems of indirect control [15]. In this sense, the problem is not only whether a user's post is removed, but also whether the architecture of visibility, recommendation, and monetisation has a chilling effect on public discourse.

The algorithmic dimension of moderation has been further developed by Gorwa, Binns, and Katzenbach, who describe algorithmic content moderation as a socio-technical system involving automated classification, human review, platform policy, and institutional incentives [16]. This understanding is essential for legal analysis because errors in moderation are not accidental defects of isolated decisions; they may be produced by the design of the system itself. When a classifier fails to distinguish documentation of war crimes from prohibited graphic violence, or satire from incitement, the resulting restriction can affect not only individual expression but also collective memory, journalism, and evidence preservation.

Recent research under the Digital Services Act [17] has brought this debate closer to positive law. Golunova argues that the DSA creates important transparency and accountability mechanisms, but that its tools must still be evaluated against the specific risks of algorithmised content moderation [5]. De Gregorio places European platform regulation within a broader constitutional transformation in which public law increasingly responds to private power in the digital sphere [18]. These studies are valuable for Ukraine because European integration requires not only the formal approximation of legislation, but also careful adaptation of European procedural safeguards to Ukrainian security conditions.

The literature, however, does not fully resolve the Ukrainian question. Most existing studies focus either on the European Union and the United States, or on general platform governance without considering the conditions of a state defending itself against a large-scale information and military attack. The Ukrainian case requires a model that protects the user from arbitrary private moderation, protects society from hostile manipulation and simultaneously prevents the state from obtaining unlimited power over online speech. This gap determines the specific contribution of the present article.

Materials and Methods

The article is based on qualitative legal research and combines doctrinal, comparative legal, functional, and systemic methods. The doctrinal method was used to analyse the legal content of constitutional and international guarantees of freedom of expression, the legal status of online platforms, and the procedural requirements applicable to moderation decisions. This method made it possible to determine whether existing legal categories, such as intermediary, editor, publisher, host, or media service provider, are sufficient to describe the role of very large online platforms in the contemporary information environment.

The comparative legal method was applied to four regulatory models: the European Union, the United States, the United Kingdom, and Taiwan. These jurisdictions were selected according to functional relevance rather than geographical similarity. The European Union was selected because the Digital Services Act is currently the most developed procedural model of platform accountability and is especially relevant for Ukraine in light of European integration [17]. The United States was selected because Section 230 of the Communications Decency Act remains the paradigmatic example of broad intermediary immunity and market-oriented platform regulation [19]. The United Kingdom was selected because the Online Safety Act 2023 represents a preventive and risk-based approach to service design and online harms [20; 21]. Taiwan was selected because it demonstrates an alternative model of democratic resilience based on civic technology, rapid response, public communication, and limited reliance on punitive takedown mechanisms [22; 23].

The functional method was used to compare not only formal legal rules, but also the functions performed by each model. The analysis, therefore, asks what each model actually does: whether it provides notice and reasons to the user; whether it creates an effective appeal mechanism; whether it requires independent review; whether it assesses systemic risks generated by platform architecture; whether it creates a channel of cooperation between the state, platforms, and civil society; and whether it avoids censorship. This approach is necessary because the same legal term may perform different functions in different regulatory systems.

The systemic method was used to evaluate algorithmic moderation as part of a broader socio-technical and legal system. The study does not treat a single deletion of content or account blocking as an isolated event. It analyses moderation as a chain of rules, automated detection, human review, platform incentives, appeal procedures, transparency reports, regulator competence and user remedies. This makes it possible to distinguish individual error from systemic risk. The method is particularly important in the Ukrainian context, where an erroneous platform decision may affect not only private communication but also journalism, evidence of international crimes, volunteer networks, and national security.

The empirical basis of the study is limited to documented public materials, including reports by civil society and international organisations concerning the impact of platform moderation on Ukrainian media and users [7; 8]. The article does not conduct an independent quantitative measurement of moderation decisions. This limitation is justified by the lack of full access to platform datasets and by the legal nature of the research. The purpose is not to calculate the total number of wrongful restrictions, but to

develop a normative and institutional model capable of responding to such restrictions when they occur. Accordingly, the conclusions are formulated as legal and policy proposals rather than statistical findings.

The research criteria were the following: protection of freedom of expression; availability of reasons and transparency; accessibility of appeal; independence of review; prevention of systemic risks; compatibility with national security; and resistance to state censorship. These criteria were applied to each foreign model and then used to design a hybrid regulatory proposal for Ukraine.

Results and Discussion

From state interference to private platform governance

The main result of the analysis is that freedom of expression in the digital environment can no longer be understood only as protection against state interference. In practice, the ability of a person, journalist, public authority, or civil society organisation to participate in public debate depends on the decisions of platforms whose rules are global, standardised, and often insufficiently sensitive to local context. A post may be removed, its visibility may be reduced, the account may be suspended, advertising may be disabled, or the content may be deprived of algorithmic distribution. Each of these measures may be formally private and contractual, but its social effect can resemble a public restriction on expression.

This does not mean that platforms should be deprived of moderation powers. On the contrary, moderation is necessary for removing unlawful content, reducing coordinated manipulation, protecting children, and preventing incitement to violence. The problem lies in the lack of procedural guarantees. When a platform restricts content without explaining the exact rule, without indicating whether the decision was automated, without providing an effective appeal, and without considering linguistic or wartime context, the user is placed in a jurisdictional vacuum. The user is formally bound by private rules, while the state is often unable to provide a practical remedy.

For Ukraine, this vacuum has a security dimension. The documentation of attacks on civilians, images of destruction, reports from occupied territories, and fundraising communications may be misclassified by automated systems. The CEDEM report demonstrates that Ukrainian media and content creators had to adapt their communication strategies because of blocks and insufficient communication from platforms; a significant number of respondents resorted to content adaptation or self-censorship to avoid further restrictions [7]. This is not simply a private inconvenience. If the fear of platform sanctions leads media actors to avoid publishing

evidence of war crimes, the effect reaches public memory, accountability, and international awareness.

At the same time, the Ukrainian state cannot solve the problem by taking direct control over permissible speech. Such an approach would contradict the constitutional logic of freedom of expression and would be vulnerable to abuse. Therefore, the proper legal response should be procedural rather than censorial. The state should not decide the truth of every disputed statement. It should establish guarantees that moderation decisions affecting users in Ukraine are reasoned, contestable, reviewable, and sensitive to the local legal and security context.

The European Union: procedural accountability and systemic risks

The Digital Services Act is the most useful comparative model for Ukraine because it shifts platform regulation from substantive state control over speech to procedural accountability. The DSA requires clearer terms and conditions, notice-and-action mechanisms, statements of reasons for restrictions, internal complaint-handling systems, out-of-court dispute settlement, and specific obligations for very large online platforms and search engines [17]. Its logic is not that the state should decide every moderation dispute. Rather, platforms must organise their decision-making in a way that is transparent, accountable, and open to challenge.

Several elements are especially important. Article 17 of the DSA requires online platforms to provide statements of reasons when they impose certain restrictions. Article 20 establishes an internal complaint-handling system. Article 21 creates the possibility of out-of-court dispute settlement. Articles 34 and 35 require very large online platforms and very large online search engines to assess and mitigate systemic risks, including risks related to fundamental rights, civic discourse, electoral processes, and public security [17]. These rules directly respond to the weaknesses of algorithmic moderation: opacity, lack of remedy, and insufficient attention to systemic effects.

The European model is not perfect. Golunova notes that the DSA creates innovative mechanisms, but that many of them are not specifically tailored to the peculiar threats posed by algorithmic content moderation [5]. The model also depends on institutional capacity, regulator independence, access to platform data, and effective certification of dispute settlement bodies. For Ukraine, these challenges would be even more serious because enforcement against transnational platforms requires both legal authority and political leverage.

Nevertheless, the DSA provides a normative core for Ukraine. Its strongest lesson is that protection of freedom of expression can be strengthened without transforming the state into a content censor. Procedural obligations,

risk assessment, and independent review are compatible with constitutional democracy because they regulate decision-making architecture rather than impose a single official truth.

The United States: immunity and market logic

The United States model is based on a different assumption. Section 230 of the Communications Decency Act provides that an interactive computer service shall not be treated as the publisher or speaker of information provided by another information content provider, and it also protects good-faith restrictions of objectionable material [19]. Historically, this rule supported innovation, protected platforms from excessive liability, and enabled the development of the open internet. It also reflects a strong distrust of state regulation of speech.

For the present study, the American model is useful primarily as a warning. Broad immunity can protect platforms from being overburdened by liability, but it does not automatically create effective remedies for users. If a platform wrongly removes content, reduces visibility, or suspends an account, the affected person may have very limited practical options. Market competition does not solve the problem when several platforms control the main channels of public communication. Nor does contractual consent solve the problem where the user has no realistic possibility to negotiate the terms of service.

The American model, therefore, cannot be transferred to Ukraine as a complete solution. In a wartime information environment, the absence of procedural obligations would leave Ukrainian users, journalists, and civil society organisations dependent on discretionary corporate decisions. At the same time, the American experience is valuable because it warns against overcorrection. A Ukrainian model should not impose such liability on platforms that they remove lawful but controversial speech out of fear of sanctions. The appropriate balance lies between immunity without accountability and state pressure without freedom.

The United Kingdom: safety by design and risk-oriented duties

The Online Safety Act 2023 represents a preventive and risk-based regulatory model. Instead of focusing only on individual items of content, it imposes duties related to the systems and processes through which online services manage illegal and harmful content, especially in relation to children and other vulnerable users [20]. Ofcom explains its role under the Act as ensuring that companies have effective systems in place to protect users from harm [21]. This model is important because it treats online harm as something produced not only by individual users, but also by the design of the service.

For Ukraine, the safety-by-design logic is relevant in the national security context. Recommender systems, viral amplification tools, anonymous channels, livestreaming, paid promotion, and automated account networks can be used not only for ordinary expression but also for coordinated information operations. A platform's architecture may enable manipulation even when no single piece of content is obviously unlawful. Therefore, a Ukrainian law should require large platforms to assess whether their functions create foreseeable risks to national security, democratic processes, access to reliable information and the safety of civilians during war.

However, the British model also requires caution. A broad online safety framework may create incentives for platforms to remove content defensively. If safety obligations are vague, they may reduce freedom of expression, especially for controversial political speech. For Ukraine, which faces both external information aggression and internal risks of excessive restriction, safety by design must be combined with strong freedom-of-expression safeguards. Risk assessment should not become a justification for suppressing legitimate criticism, journalism, satire, or documentation of violence.

Taiwan: social resilience and non-punitive coordination

Taiwan offers a different lesson. Its approach to disinformation has often relied on digital resilience, civic technology, rapid clarification, and cooperation between public institutions and civil society, rather than on mass takedown powers. The Ministry of Digital Affairs describes Taiwan's response to disinformation as one in which civil society took the lead in quarantining falsehoods and paying attention to information manipulation, enabling Taiwan to combat the infodemic with no takedowns [22]. Cofacts, a civic fact-checking project integrated with LINE, functions as a crowdsourced system connecting suspicious messages with fact-checking reports and alternative explanations [23].

This model is particularly attractive for Ukraine because it separates response to disinformation from censorship. The state does not need to block every false statement in order to reduce the harm of manipulation. It can support rapid verification, public explanation, media literacy, open data, and communication channels with platforms. In wartime, speed is essential. A false claim about mobilisation, evacuation, attacks on civilians, or international support may cause harm long before a court or regulator reaches a formal decision. A resilience model, therefore, complements, rather than replaces, legal remedies.

The weakness of the Taiwanese approach is that it cannot by itself compel global platforms to correct wrongful moderation or redesign harmful systems. Civic resilience is effective against falsehoods, but it does not

solve the problem of arbitrary private restrictions. Ukraine, therefore, needs Taiwan's speed and trust, but also the European procedural architecture and a risk-based approach to platform systems.

Ukrainian legal gaps

The Ukrainian framework contains important general rules but lacks a special platform governance regime. The Law of Ukraine on Information defines general principles of information relations and access to information [9]. The Law of Ukraine On Media modernises media regulation and addresses aspects of media activity, audiovisual services, and the powers of the national regulator [10]. The Law of Ukraine on Personal Data Protection is relevant to the processing of user data in digital environments [11]. However, none of these acts creates a full procedural framework for the moderation decisions of very large online platforms.

Several gaps are especially visible. First, Ukrainian users do not have a statutory right to receive a detailed and standardised explanation of why content was removed, visibility was restricted, or an account was blocked. Secondly, there is no specialised independent mechanism for reviewing platform decisions after the internal appeal has failed. Thirdly, Ukrainian law does not require large platforms to assess systemic risks caused by their recommendation systems, content moderation systems, advertising architecture, or design choices in the Ukrainian context. Fourthly, there is no permanent institutional channel through which Ukrainian authorities, civil society, and platforms can address repeated false positives, wartime context errors, and coordinated disinformation campaigns without resorting to direct censorship.

The absence of such a regime creates an asymmetry of power. A Ukrainian journalist, volunteer, or media outlet may depend on a platform for reach, fundraising, or international visibility, but the platform may treat the Ukrainian case as one among millions of global moderation events. The domestic court may not be an effective remedy when the platform has no meaningful local presence or when the dispute is resolved internally by automated systems. This asymmetry justifies legislative intervention, but only if the intervention is designed as procedural accountability rather than control over viewpoints.

A hybrid model of platform governance for Ukraine

The most appropriate solution for Ukraine is a hybrid model of platform governance. Its first component should be procedural accountability. A special law on digital services and platform governance should require large platforms operating in Ukraine to provide a reasoned notification for every significant moderation decision affecting content visibility, account

status, monetisation, or access to services. The notification should include the specific rule applied, the factual basis of the decision, the type of measure, the duration of restriction, whether automated tools were used, and information on available appeal mechanisms. This would adapt the procedural logic of the DSA to the Ukrainian context.

The second component should be an internal and independent appeal. Platforms should be required to maintain accessible internal complaint mechanisms in Ukrainian and, where relevant, English. After exhaustion of internal appeal, users should have access to an independent national or certified extrajudicial review mechanism. This could be organised as a Digital Ombudsman or a specialised board under an independent regulator. The body should not have general censorship powers. Its competence should be limited to assessing whether the platform decision is reasoned, proportionate, consistent with Ukrainian law and international human rights standards, and whether the platform has properly considered the Ukrainian context.

The third component should be systemic risk assessment. Very large platforms with a significant number of users in Ukraine should periodically assess risks arising from their algorithmic systems, content moderation policies, recommender systems, advertising tools, and crisis-response procedures. Such an assessment should cover risks to freedom of expression, access to reliable information, democratic processes, national security, children, journalists, human rights defenders, and users in occupied or frontline territories. The assessment should not be limited to illegal content. It should also address lawful but harmful manipulation, coordinated inauthentic behaviour, and wrongful suppression of socially significant lawful content.

The fourth component should be a Digital Resilience Coordination Centre. Such a centre could operate under the Ministry of Digital Transformation with participation of civil society, independent fact-checkers, media organisations, and academic experts. It should not be authorised to order content blocking. Its tasks should be rapid verification, communication with platforms, collection of recurring moderation problems, preparation of contextual guidance for platforms, and public explanation during disinformation campaigns. In this respect, Ukraine can borrow from the Taiwanese model without importing censorship powers.

The fifth component should be proportional enforcement. Sanctions should apply only where a platform systematically refuses to cooperate with legally established procedural mechanisms, ignores binding review decisions, fails to provide required transparency, or repeatedly neglects systemic risks affecting Ukrainian users. Sanctions should be economic and institutional

rather than viewpoint-based. The object of enforcement should be non-compliance with procedure, transparency, or risk-management duties, not disagreement with a particular editorial position.

Such a hybrid model would restore the role of the state as guarantor of freedom of expression without transforming it into the arbiter of truth. It would also recognise that digital sovereignty today means not isolation from global platforms, but the ability of a democratic legal order to require minimum standards of fairness, transparency, and accountability from private actors that structure public discourse.

Conclusions

The study demonstrates that freedom of expression in the digital environment is increasingly affected by private algorithmic governance rather than only by direct state interference. Global platforms define the visibility, reach, and practical effectiveness of speech through rules, classifiers, recommender systems, and enforcement practices. For Ukraine, this creates a specific constitutional and security problem: wrongful moderation may restrict journalism, documentation of war crimes, volunteer communication, and public mobilisation, while excessive state control over platforms would threaten censorship.

The comparative analysis shows that no foreign model can be copied mechanically. The European Union provides the strongest procedural safeguards and systemic risk logic; the United States protects innovation but leaves users with weak remedies; the United Kingdom contributes a safety-by-design perspective; and Taiwan demonstrates the value of rapid, trust-based, non-punitive resilience. Ukraine should therefore develop a hybrid model combining reasoned moderation decisions, internal and independent appeal, systemic risk assessment, institutional communication with platforms and a digital resilience centre without blocking powers. Such a model would strengthen constitutional freedom of expression, create practical remedies for users and media, and improve national security in the face of continuing information aggression.

References

[1] Constitution of Ukraine. (June 28, 1996). Retrieved from https://zakon.rada.gov.ua/laws/show/254к/96-вр#Text.

[2] Council of Europe. (1950). *Convention for the Protection of Human Rights and Fundamental Freedoms*. Retrieved from https://www.echr.coe.int/documents/convention_eng.pdf.

[3] Delfi AS v. Estonia [GC], Application No. 64569/09, European Court of Human Rights. (2015). Retrieved from https://hudoc.echr.coe.int/eng?i=001-155105.

[4] Cengiz and Others v. Turkey, Applications Nos. 48226/10 and 14027/11, European Court of Human Rights. (2015). Retrieved from https://hudoc.echr.coe.int/eng?i=001-156825.

[5] Golunova, V. (2025). Algorithmic Content Moderation and Freedom of Expression under the Digital Services Act. European Review of Public Law, 37(2), 1-36.
[6] Article 19. (2023). *Content moderation and freedom of expression handbook.* Retrieved from https://www.article19.org/resources/content-moderation-freedom-of-expression-handbook/.
[7] Centre for Democracy and Rule of Law, Internews Ukraine, & Lviv Media Forum. (2023). *Social Media's Impact on the Ukrainian News and Publishing Space after the Full-Scale Invasion.* Retrieved from https://cedem.org.ua/en/research/social-media-ukraine-war/.
[8] Freedom House. (2023). *Freedom on the Net 2023: Ukraine.* Retrieved from https://freedomhouse.org/country/ukraine/freedom-net/2023.
[9] Law of Ukraine No. 2657-XII "On Information". (October 02, 1992). Retrieved from https://zakon.rada.gov.ua/laws/show/2657-12#Text.
[10] Law of Ukraine No. 2849-IX "On Media". (December 13, 2022). Retrieved from https://zakon.rada.gov.ua/laws/show/2849-20#Text.
[11] Law of Ukraine No. 2297-VI "On Personal Data Protection". (June 01, 2010). Retrieved from https://zakon.rada.gov.ua/laws/show/2297-17#Text.
[12] Celeste, E. (2019). Digital Constitutionalism: A New Systematic Theorization. *International Review of Law, Computers & Technology, 33*(1), 76-93. https://doi.org/10.1080/13600869.2019.1562604.
[13] Suzor, N. (2018). Digital Constitutionalism: Using the Rule of Law to Evaluate the Legitimacy of Governance by Platforms. *Social Media + Society, 4*(3), 1-11. https://doi.org/10.1177/2056305118787812.
[14] Gillespie, T. (2018). Regulation of and by Platforms. In J. Burgess, A. Marwick, & T. Poell (Eds.). *The SAGE Handbook of Social Media* (pp. 254-278). SAGE. https://doi.org/10.4135/9781473984066.n15.
[15] Balkin, J.M. (2018). Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation. *UC Davis Law Review, 51*(3), 1149-1210. Retrieved from https://lawreview.law.ucdavis.edu/issues/51/3/Articles/51-3_Balkin.pdf.
[16] Gorwa, R., Binns, R., & Katzenbach, C. (2020). Algorithmic Content Moderation: Technical and Political Challenges in the Automation of Platform Governance. *Big Data & Society, 7*(1), 1-15. https://doi.org/10.1177/2053951719897945.
[17] 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). (2022). *Official Journal of the European Union, L 277,* 1-102. Retrieved from https://eur-lex.europa.eu/eli/reg/2022/2065/oj.
[18] De Gregorio, G. (2022). *Digital Constitutionalism in Europe: Reframing Rights and Powers in the Algorithmic Society.* Cambridge University Press. https://doi.org/10.1017/9781009071215.
[19] Communications Act of 1934, 47 U.S. Code § 230. (2024). Retrieved from https://www.law.cornell.edu/uscode/text/47/230.
[20] Online Safety Act 2023, 50 (UK). (2023). Retrieved from https://www.legislation.gov.uk/ukpga/2023/50/contents.
[21] Ofcom. (2025). *Online Safety: Guidance and Implementation Materials.* Retrieved from https://www.ofcom.org.uk/online-safety.
[22] Ministry of Digital Affairs, Taiwan. (2022). *Introduction to Moda.* Retrieved from https://moda.gov.tw/en/.
[23] Cofacts. (2024). *Cofacts: Message Reporting Chatbot and Crowd-Sourced Fact-Checking Platform.* Retrieved from https://cofacts.tw/.

Alona D. Harahata
Ph.D. Student of the Department of Theory and History of Law
Yaroslav Mudryi National Law University
61024, 77 Hryhoriia Skovorody Str., Kharkiv, Ukraine
e-mail: a.d.garagata@nlu.edu.ua
ORCID 0009-0002-9396-6359

Альона Дмитрівна Гарагата
аспірантка кафедри теорії та історії права
Національний юридичний університет імені Ярослава Мудрого
61024, вул. Григорія Сковороди, 77, Харків, Україна
e-mail: a.d.garagata@nlu.edu.ua
ORCID 0009-0002-9396-6359

Suggested Citation: Harahata, A.D. (2026). Algorithmic Platform Moderation and Freedom of Expression in Ukraine. *Theory and Practice of Jurisprudence, 1*(29), 44-58. https://doi.org/10.21564/2225-6555.2026.29.360305.

Submitted: 17.04.2026
Revised: 20.05.2026
Approved: 21.05.2026
Published online: 28.05.2026