

# GOVERNING AI AND CONTENT MODERATION ONLINE: THE EU REGULATORY APPROACH AND GLOBAL MULTISTAKEHOLDER DYNAMICS

**Assoc. Prof. Denitza Topchiyska, PhD**

*New Bulgarian University*

## ***Abstract:***

*This article examines the key provisions of EU legislation concerning the regulation of online content moderation through the use of artificial intelligence (AI). These provisions, drawn from the General Data Protection Regulation (GDPR), the Digital Services Act (DSA), and the AI Act, are analysed in the context of the established global internet governance framework and the European Union's specific position on its development, including its foundational values and defining principles. The analysis supports the identification of key risks and challenges to the effective implementation of the EU's legislative framework and ultimately proposes recommendations to address these issues and support the consistent and rights-based application of the relevant regulations.*

**Keywords:** artificial intelligence, online content moderation, internet governance, multistakeholder approach

During the 2020s, the discourse on content moderation in social media has emerged as a crucial dimension in broader debates concerning human rights and democracy in the digital era. This discourse is fuelled by the expanding role of social media platforms in shaping public opinion and influencing democratic processes, including political elections. In June 2024, the European Commission launched the DSA Transparency Database,<sup>1</sup> mandating online platforms to report their decisions regarding content moderation, including the rationale behind them. To date, the database has registered over 10 billion notifications, half of which have been processed using automated tools or artificial intelligence. As part of its comprehensive strategy to ensure the integrity of

---

<sup>1</sup> DSA Transparency Database: <https://transparency.dsa.ec.europa.eu/?lang=en>

information in the digital sphere, known as the European Democracy Shield,<sup>2</sup> the EU is adopting specific measures to address the risks associated with the use of AI in the complex task of balancing the protection of freedom of expression with the need to counter the dissemination of illegal and harmful content online.

This article aims to examine the key provisions of EU legislation concerning the regulation of online content moderation through the use of artificial intelligence (AI). These legal provisions will be analysed in the context of the established global internet governance model and the EU's specific stance towards its development, including its foundational values and defining characteristics. The ensuing analysis will support the identification of significant risks and challenges to the effective implementation of the EU's legislative framework, leading to proposed recommendations for their mitigation.

## **1. The EU on AI and content moderation online**

Within the broader European Democracy Shield regulatory framework, issues concerning the application of AI in online content moderation addressed primarily in the General Data Protection Regulation (GDPR)<sup>3</sup>, the Digital Services Act (DSA)<sup>4</sup>, and the Artificial Intelligence Act (AI Act)<sup>5</sup>.

In 2016, the EU enacted the General Data Protection Regulation (GDPR), which became directly applicable across all Member States in 2018, thereby establishing a horizontal framework governing the automated processing of personal data. In its article 22, paragraph 1, the GDPR sets out a general prohibition on decisions based solely on automated processing, including profiling. In 2018, the Article 29 Data Protection Working Party adopted Guidelines on the application of this provision, stating that the prohibition applies regardless of whether the data subject takes any action related to the processing of their personal data.<sup>6</sup> It has been clarified that decisions „based solely“ on automated

---

<sup>2</sup> BRIEFING EPRS | European Parliamentary Research Service, Author: Naja Bentzen, December 2024: „Information integrity online and the European democracy shield“ ([https://www.europarl.europa.eu/thinktank/en/document/EPRS\\_BRI\(2024\)767153](https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2024)767153))

<sup>3</sup> Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ L 119, 4.5.2016, p. 1-88

<sup>4</sup> 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), OJ L 277, 27.10.2022, p. 1-102

<sup>5</sup> Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), OJ L, 2024/1689, 12.7.2024.

<sup>6</sup> Article 29 Data Protection Working Party, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, last Revised and Adopted on 6 February 2018

processing are to be interpreted as those where the decision-making process is entirely devoid of human intervention. Furthermore, the general prohibition on fully automated decision-making applies in cases where such decisions produce legal effects or similarly significantly affect the data subject. Thus, online content moderation processes conducted through automated means fall within the scope of this provision, as they have a direct and potentially far-reaching impact on individuals' freedom of expression.

The GDPR allows for exceptions to the general prohibition on decision-making based solely on automated processing when such processing is expressly authorized by EU law or the law of a Member State, when it is necessary for the conclusion or performance of a contract, or when the data subject has given their explicit consent. In such cases, the Regulation requires data controllers to implement safeguards, including the right to information as detailed in Articles 13 and 14. Specifically, they must provide meaningful information regarding the logic involved, along with the significance and anticipated consequences of such processing for the data subject. Additionally, the data controller shall establish appropriate safeguards, encompassing the right to human intervention and the right to contest the decision.<sup>7</sup>

In 2022, the EU adopted the Digital Services Act (DSA), establishing a novel legal regulation for content moderation on online platforms to counter the spread of illegal and harmful content, including disinformation, misinformation and propaganda. Simultaneously, the previously adopted self-regulatory mechanisms, the EU Code of Practice on Disinformation<sup>8</sup> and the EU Code of Conduct on Countering Illegal Hate Speech Online,<sup>9</sup> continued to be applied. In early 2025, the European Commission and the European Board for Digital Services formally endorsed the integration of these documents as Codes of Conduct within the framework of the DSA to facilitate enhanced oversight by EU institutions over their implementation.

The DSA sets out obligations for online platforms regarding content moderation but does not impose a general duty on social media platforms to monitor or proactively search for illegal content shared by their users. The Regulation requires online platforms to implement notice-and-action mechanisms that allow users to notify them of specific information hosted on their service which the notifier considers to be illegal. These systems must be easily accessible and permit the electronic submission of notifications. The DSA sets standards for the processing of such notifications, stipulating that platforms must act promptly, in good faith, and in an impartial and objective manner. Platforms are also required

---

<sup>7</sup> Article 22, paragraph 3 of GDPR

<sup>8</sup> The 2022 Code of Practice on Disinformation: <https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation>

<sup>9</sup> The EU Code of conduct on countering illegal hate speech online: [https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online\\_en#theeucodeofconduct](https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online_en#theeucodeofconduct)

to inform the notifier of the decision taken, indicate whether automated means were used in the process, and outline the available legal remedies. Additionally, online platforms must provide each affected service recipient with a clear and specific explanation of the reasons for any imposed access restrictions to content deemed illegal or in violation of their terms of service. Such measures may include restricting access to content, temporary suspension, or account termination. The explanation must be clear, easily understandable, and as accurate and specific as possible, reflecting the particular circumstances of the case.

In regulating content moderation processes, the DSA introduces safeguards to ensure that such processes are not carried out solely by automated means. Social media platforms are required to establish an easily accessible internal complaint-handling system for reviewing content moderation decisions. These decisions must be made under the supervision of staff with appropriate qualifications, rather than relying exclusively on automated tools. In addition to the legal avenues available for challenging platform decisions, the DSA also provides for the establishment of alternative dispute resolution bodies, certified by the Digital Services Coordinators. These bodies may be created or supported by the state. When a user files a complaint with such a body, the platform is not permitted to refuse participation in the dispute resolution process.

Through the DSA, the European Union seeks to establish clear standards for procedures implemented by private online platforms, ensuring that they conform to the principles of legality applicable to public regulation. The objective is to address the structural shortcomings of self-regulation by setting enforceable requirements for transparency and procedural fairness in decision-making carried out by private entities. It is of particular importance that the DSA reaffirms the obligation to respect and uphold core foundational principles such as the rule of law, accountability, and the protection of fundamental rights, even when regulatory or governance functions are exercised by private actors.

In 2024, the European Union adopted the Artificial Intelligence Act (AI Act), which shall enter into full application in August 2026. The objective of the Regulation is to promote the development and uptake of human-centric and trustworthy AI, while ensuring the protection of fundamental rights, democracy, and the principles of the rule of law against the potentially harmful effects of AI systems within the Union. In cases where online platforms, including social media employ automated content moderation systems that fall within the scope of the term „AI system,“<sup>10</sup> they are required to comply with the obligations set forth by the AI Act. The Regulation addresses risks associated with AI, such as bias, discrimination, and shortcomings in accountability, categorizing them into four levels of risk according to the specific use of AI. Depending on the characteristics of the deployed automated system and the

---

<sup>10</sup> Art. 3, par. 1 (1) of the AI Act

specific risks it entails, the Regulation stipulates the application of concrete rules, including, *inter alia*, obligations of transparency.

The AI Act also refers to the 2019 Ethics Guidelines for Trustworthy Artificial Intelligence,<sup>11</sup> formulated by the High-Level Expert Group on AI (AI HLEG), which was appointed by the European Commission.<sup>12</sup> Although these guidelines do not have legally binding force, they are regarded as complementary to the mandatory requirements set out in this Regulation, fostering the development of coherent, trustworthy, and human-centric AI systems in conformity with the Charter and the foundational values of the Union.<sup>13</sup> The AI HLEG has articulated seven non-binding ethical principles designed to ensure that AI is both trustworthy and ethically sound: human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity, non-discrimination and fairness; societal and environmental well-being; and accountability. Where feasible, these principles ought to be incorporated into the design and deployment of AI systems. The AI Act stipulates that these principles shall be taken into consideration in the development of codes of conduct under the regulation. All stakeholders, including industry, academia, civil society, and standardisation bodies, are encouraged to integrate these ethical principles, as appropriate, into the development of voluntary best practices and standards.

With the adoption of the GDPR, the DSA, and the AI Act, the EU has embraced a comprehensive approach to establishing a „co-regulatory backstop.“ This framework supplements self-regulatory mechanisms with legal safeguards, ensuring both transparency and accountability from digital platforms towards regulators and users, while simultaneously aiming to prevent content censorship. The adopted regulations provide for soft law tools, such as guidelines and best practices, intended to specify and elaborate upon the general rules of the legislative provisions, thereby facilitating their implementation by the relevant actors. In the development of these instruments, the expertise of private actors is frequently relied upon, ensuring that the resulting frameworks are both practically applicable and informed by sector-specific knowledge. The approach adopted by the EU requires each Member State to designate independent national authorities responsible for supervising the implementation of the regulatory framework. These authorities are also entrusted with facilitating the joint development and application of co-regulatory instruments, in close collaboration with all relevant stakeholders. The aim of these instruments is to combine the legal certainty and enforceability of hard law with the flexibility and adaptability of soft law mechanisms.

---

<sup>11</sup> Ethics Guidelines for Trustworthy Artificial Intelligence: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

<sup>12</sup> Recital 27 of the AI Act

<sup>13</sup> European Union. (2012). *Charter of Fundamental Rights of the European Union*. Official Journal of the European Union, C 326, 391-407.

The EU's regulatory approach seeks to provide more robust safeguards for users, ensuring that content moderation procedures are transparent, well-justified, and inclusive of users' perspectives in the decision-making process. Simultaneously, the role of national governments is being reinforced through the involvement of competent public authorities in the oversight of social media platforms, including their content moderation practices. While this model is consistent with Europe's political and legal tradition, it may appear less intuitive in more libertarian regulatory environments, such as that of the United States. However, to fully assess the effectiveness and coherence of the EU's approach, it is essential to consider it within the broader context of global internet governance – particularly the multistakeholder model, which underpins the development of the Internet as an open and global network. Clarifying the EU's position within this global governance landscape remains crucial to ensuring both normative consistency and international legitimacy.

## 2. The global internet governance model

The concept of governance is central to contemporary regulatory theory, where it reflects a modern understanding of regulation as a flexible and evolving system of norms, institutions, and practices.<sup>14</sup> Rather than a static set of rules, regulation is seen as a dynamic response to the prevailing social consensus, shaped by changing societal expectations and conditions. This dynamic nature is evident not only in the substantive content of regulation but also in the instruments employed, which must adapt to ensure effective and legitimate governance in a transforming environment.

The modern paradigm of governance refers to a broad concept of regulation that encompasses both public and private normative frameworks, distinguishing it from the traditional notion of regulation as a purely governmental function. In the context of internet regulation, the term governance has consistently been used to reflect the complex and multilayered process of developing and establishing mechanisms for the coordinated management of the global network. Within this framework, governance refers to the recognition and implementation of collectively accepted rules and procedures, whether initiated by public institutions or private actors, and regardless of whether they emerge through top-down authority or bottom-up, negotiated processes.<sup>15</sup> Governance operates across multiple levels of social organization, ranging from internal organizational settings to national and global contexts. It may be institutionalized through a variety of mechanisms, including formal laws, regulatory frame-

---

<sup>14</sup> „Modern regulation is the set of norms, institutions, and practices that guarantees the stability of expectations.“ - Santos, B. de S. (2020). *Toward a new legal common sense law, globalization, and emancipation* (Third edition). Cambridge University Press, p. 2

<sup>15</sup> William J. Drake (2004) Reframing Internet Governance Discourse: Fifteen Baseline Propositions. - In, Don MacLean, ed., *Internet Governance: A Grand Collaboration*. New York: The United Nations Information and Communication Technology Taskforce, pp. 122-161. (p. 125)

works, public policies, as well as decisions and procedures originating from non-state actors.

In 2005, the Working Group on Internet Governance (WGIG) defined Internet governance as „the development and application by governments, the private sector and civil society, in their respective roles, of shared principles, norms, rules, decision-making procedures, and programmes that shape the evolution and use of the Internet“.<sup>16</sup> The broad definition adopted by the Working Group is regarded as a pivotal step in the development of Internet governance, as it enables the multi-stakeholder approach to become established and to extend into a wider range of areas. The broad definition adopted by the Working Group on Internet Governance (WGIG) constitutes a landmark in the conceptual evolution of Internet governance, as it provides the normative foundation for the institutionalisation and expansion of the multi-stakeholder model across a wide range of regulatory domains. This model is widely regarded as paradigmatic for Internet governance, reflecting its decentralised, inclusive, and participatory character. The WGIG, composed of 40 members drawn from governments, the private sector, and civil society, reached its conclusions on the basis of equal participation and deliberation. Members acted in their personal capacity rather than as representatives of institutional interests, thereby affirming the legitimacy of the multi-stakeholder approach as both a procedural principle and a normative standard for global Internet governance.

The WGIG’s definition underscores the inclusive character of Internet governance by affirming the involvement of governments, the private sector, and civil society, while recognising their distinct roles, interests, and levels of engagement across policy domains. It extends beyond the technical management of names and addresses by ICANN to encompass broader public policy issues, including the governance of critical Internet resources, cybersecurity, and the developmental dimensions of Internet access and use.

The adoption of the broad definition of Internet governance continues the legacy of the Internet’s early architects, who deliberately embedded its core features into the network’s design. Although the Internet originated as a government-funded project in the United States, its development was primarily driven by academic and technical communities. With the formal opening of the Internet to commercial use in 1991, the U.S. government embraced a light-touch regulatory model, which facilitated the emergence of a bottom-up, self-regulating technical community that played a foundational role in shaping the Internet’s institutional and normative architecture.<sup>17</sup>

At the core of the internet evolution lay the influence of cyber-libertarian thought, which advanced the view that the Internet should remain a domain

---

<sup>16</sup> Working Group on Internet Governance. (2005). *Report of the Working Group on Internet Governance*. United Nations. <https://www.wgig.org/docs/WGIGREPORT.pdf>

<sup>17</sup> U.S. Department of Commerce. (1997). *A framework for global electronic commerce*.



of individual freedom, shielded from excessive state control. This ideological framework, rooted in scepticism toward the capacities and legitimacy of traditional governmental institutions, contributed decisively to the adoption of a distributed, non-hierarchical model of governance premised on community-driven innovation and voluntary coordination. The Internet's technical architecture was intentionally designed to promote decentralisation, resilience, and autonomy among its participants. Its distributed structure precludes centralised control and reflects core design principles that prioritise adaptability and the independence of individual network entities. Foundational concepts such as network neutrality, openness, and the end-to-end principle are embedded within this architecture. These principles ensure non-discriminatory data transmission, facilitate global interoperability and information exchange, and allocate complex functions to the network's endpoints – thereby fostering innovation and maintaining the Internet as a universal, open communication space.

Internet governance functions within a highly dynamic and decentralized network of actors, procedures, and institutions, which inherently resists systematic organization and poses considerable challenges to achieving coherent and coordinated action across its diverse components.<sup>18</sup> From a systematic standpoint, the Internet Corporation for Assigned Names and Numbers (ICANN) proposes a three-layer framework for Internet governance, comprising the infrastructure layer, the logical layer, and the economic and societal layer.<sup>19</sup> Each governance layer presents distinct challenges that require differentiated responses, including technical standards, policy instruments, best practices, and institutional mechanisms. These responses are typically developed through collaborative multistakeholder processes involving governments, private sector actors, civil society, academia, and technical experts. The infrastructure layer encompasses the physical components of the Internet, such as cables, satellites, and exchange points, and involves national authorities, private operators, and technical specialists. The logical layer governs domain names, IP addresses, and protocols essential for the Internet's stability and interoperability, primarily through expert-driven multistakeholder frameworks. The social and economic layer addresses broader legal, cultural, and economic issues relating to Internet use and engages a diverse range of actors including states, intergovernmental organizations, private platforms, and civil society. Collectively, these layers illustrate the inherent complexity and multistakeholder nature of Internet governance. Key institutions and forums such as the Internet Corporation for Assigned Names and Numbers (ICANN), the Internet Engineering Task Force (IETF), the Internet Governance Forum (IGF), and the International Telecommunication Union (ITU) play central roles in shaping policies and coordinating governance activities across these layers.

---

<sup>18</sup> Jeanette Hofmann, *Internet Governance: A Regulative Idea in Flux*, in Ravi Kumar Jain Bandamutha (Ed.), *Internet Governance: An Introduction* (Icfai University Press, 2007), pp. 74-108.

<sup>19</sup> ICANN. (n.d.). *The Internet ecosystem* [Infographic]. ICANN. <https://www.icann.org/news/multimedia/1563>



### 3. The EU position on internet governance and the European Democracy Shield

Since the global reach of the Internet as a communications infrastructure, its governance has been the subject of sustained debate concerning the appropriate institutional model, in particular whether it should be exercised through an intergovernmental framework, led primarily by states, or through a multistakeholder approach that ensures the equal participation of governments, the private sector, civil society, and the technical and academic communities. Within this discourse, the European Union has consistently taken the position of a strong proponent of the multi-stakeholder approach, advocating for inclusive, transparent, and collaborative mechanisms in the formulation of internet governance principles, rules, and procedures. The EU remains a staunch advocate of a single, open, free, neutral and unfragmented internet, conceived as a decentralised network of networks. This vision stands in clear contrast to more centralised and state-controlled models promoted by certain governments, where access to information is restricted and user activity is subject to systematic surveillance. At the same time, the EU acknowledges that the integrity and openness of the internet can also be compromised by private actors, especially when companies establish proprietary infrastructure or enforce exclusive technical standards that may result in the fragmentation of the global internet ecosystem.<sup>20</sup>

In 2022, the European Union adopted the European Declaration on Digital Rights and Principles, a political document intended to steer the digital transformation in alignment with core European values such as digital sovereignty, democracy, and the protection of fundamental rights and freedoms. Although not legally binding, the Declaration functions as a guiding framework for the formulation of sustainable, human-centric digital policies and reinforces the principles of the Charter of Fundamental Rights of the EU by affirming the continuity of rights in both online and offline environments. It carries normative and interpretative weight, with the potential to influence EU legislation, judicial interpretation, and national digital strategies. Furthermore, the Declaration aspires to serve as a global reference point for digital rights, with the European Commission assessing progress through the annual State of the Digital Decade report.<sup>21</sup>

---

<sup>20</sup> Niestadt, M. (2024, November 29). *Internet governance: Keeping the internet open, free and unfragmented* (EPRS Briefing No. 766272). European Parliamentary Research Service. Available at: [https://www.europarl.europa.eu/thinktank/en/document/EPRS\\_BRI\(2024\)766272](https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2024)766272)

<sup>21</sup> European Commission. (2022). *European Declaration on Digital Rights and Principles for the Digital Decade*. Publications Office of the European Union. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022DC0028> /European Commission. (2022, December 14). *Commission proposes European Declaration on Digital Rights and Principles* (Press Release No. IP/22/452). European Commission - Press Corner. Available at: [https://ec.europa.eu/commission/presscorner/detail/en/ip\\_22\\_452](https://ec.europa.eu/commission/presscorner/detail/en/ip_22_452)

In line with its vision for the digital transformation of the Union, in 2022, the EU, together with 70 other like-minded states, launched the Declaration for the Future of the Internet.<sup>22</sup> The document constitutes a non-binding political commitment among its signatory partners to promote a shared, affirmative vision for the development and governance of the internet and digital technologies in the 21st century. It reasserts support for a single, open, global, and interoperable internet that upholds human rights, fosters fair competition, protects privacy, and ensures democratic accountability. The Declaration sets out a number of foundational principles, including the protection of fundamental rights and freedoms, the promotion of the free flow of information, universal and affordable connectivity, trust in the digital ecosystem (particularly through robust privacy protections), and the preservation of the multi-stakeholder model of internet governance. Participating states commit to advancing these principles globally, while acknowledging and respecting each other's regulatory autonomy within their respective jurisdictions and in accordance with both domestic legal frameworks and international legal obligations.

The EU has identified critical risks to democratic information ecosystems in the digital sphere. The online environment, as the main forum for information and expression, has become a geostrategic space where authoritarian states manipulate public discourse and deepen divisions, threatening democratic governance. This is intensified by the intersection of geopolitical rivalry and corporate competition, especially with emerging technologies like artificial intelligence that alter information flows. The EU highlights that AI-driven disinformation campaigns significantly increase these threats, undermining freedom of expression and human rights. In response, the EU calls for upholding democratic values online and advancing coordinated multilateral actions to strengthen information ecosystem resilience. These aims are central to initiatives such as the European Democracy Shield, which offers a legislative and strategic framework to protect information integrity.<sup>23</sup>

Information integrity has emerged as a central concept in international and multilateral efforts to address the complex challenges facing the digital information environment. While no single definition prevails, the concept generally reflects a positive, rights-based approach aimed at promoting access to trustworthy information, safeguarding freedom of expression, and ensuring the sustainability of the information ecosystem. In the European Union, information integrity is a key pillar of the forthcoming European Democracy Shield, launched in 2024, which integrates existing initiatives to counter foreign information

---

<sup>22</sup> European Commission & Council of the European Union. (2022, April 25). *Declaration for the Future of the Internet*. Publications Office of the European Union. Available at: <https://digital-strategy.ec.europa.eu/en/library/declaration-future-internet>

<sup>23</sup> Bentzen, N. (2024, December 10). *Information integrity online and the European democracy shield* (EPRS Briefing No. 767153). European Parliamentary Research Service. Available at: [https://www.euparl.europa.eu/thinktank/en/document/EPRS\\_BRI\(2024\)767153](https://www.euparl.europa.eu/thinktank/en/document/EPRS_BRI(2024)767153)

manipulation and interference (FIMI) with major legislative instruments such as the Digital Services Act, the AI Act, the European Media Freedom Act, and the Regulation on Transparency and Targeting of Political Advertising. These measures collectively seek to enhance the resilience and integrity of the EU's information space by combining regulation, oversight, and multistakeholder cooperation.

The European Democracy Shield is part of a wider international effort to enhance information integrity and promote responsible digital governance. This global momentum was highlighted on 22 September 2024, when world leaders at the United Nations Summit of the Future adopted the Pact for the Future, including the Global Digital Compact. The Compact provides a comprehensive framework for international digital cooperation, focusing on the governance of emerging technologies such as artificial intelligence. It calls on digital companies and social media platforms to improve transparency and accountability in key areas, including terms of service, content moderation, recommendation algorithms, and personal data processing, especially in local languages. These measures aim to empower users to make informed choices and support sustainable development and digital inclusion.<sup>24</sup>

## Conclusions

This article aimed to critically examine the key provisions of European Union legislation governing online content moderation through the use of artificial intelligence included in the GDPR, DSA and AI Act. These legal instruments were analysed in light of the global Internet governance framework and the EU's normative vision for a single, open, global, and interoperable Internet. Based on this analysis, several important conclusions may be drawn regarding the proper interpretation and effective implementation of these regulatory frameworks.

Although EU regulations such as the GDPR, DSA and AI Act are directly applicable and possess primacy over conflicting national legislation, their practical effectiveness ultimately depends on the extent to which they are successfully integrated into domestic legal systems. The capacity of national regulatory authorities to monitor compliance, enforce obligations, and engage with stakeholders is indispensable for the achievement of the Union's policy objectives. For this reason, the interpretation and application of these legal instruments at national level must be situated not only within the framework of the EU's internal policy agenda but also in light of the evolving global landscape of internet governance. International instruments and initiatives such as the United Nations' Global Digital Compact, the OECD's recommendations on AI, and the Council of Europe's work on algorithmic systems

---

<sup>24</sup> United Nations. (2024, September 22). *Global Digital Compact*. Office of the Secretary-General's Envoy on Technology. Available at <https://www.un.org/digital-emerging-technologies/global-digital-compact>

illustrate a growing convergence around common principles, including transparency, accountability, and the protection of fundamental rights. Aligning EU implementation practices with these global efforts is essential to fostering regulatory coherence, enhancing cross-border interoperability, and reinforcing the EU's role in shaping a rights-based, inclusive digital order.

Furthermore, the interpretation and application of these regulations must be anchored within the broader framework of the European Union's strategic policy on Internet governance. This policy envisions a single, open, global, and interoperable Internet, which respects fundamental rights, guarantees fair competition, ensures privacy, and reinforces democratic accountability. Therefore, any interpretative approach that compromises these foundational principles would not only conflict with the Union's normative vision but would also constitute a misapplication of the legal framework.

In addition, the implementation of EU regulations should reflect the Union's firm commitment to the multistakeholder approach in internet governance. This entails the inclusive and collaborative participation of all relevant actors, including governmental institutions, private sector organisations, civil society, technical experts and academic institutions, throughout all phases of regulatory execution. Furthermore, the effective use of soft law instruments such as guidelines, self-regulatory codes and voluntary standards should be actively encouraged in order to support and complement the binding legal framework.

Finally, the inherently dynamic and multilayered nature of Internet governance necessitates that regulatory enforcement remains attuned to the evolving institutional landscape. As technological, geopolitical, and societal developments reshape the digital environment, the implementation of EU law must remain flexible, adaptive, and responsive to ensure coherence with broader governance processes at the international, European, and national levels.

## BIBLIOGRAPHY

- Article 29 Data Protection Working Party. (2018, February 6). *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*.
- Bentzen, N. (2024, December 10). *Information integrity online and the European democracy shield* (EPRS Briefing No. 767153). European Parliamentary Research Service. [https://www.europarl.europa.eu/thinktank/en/document/EPRS\\_BRI\(2024\)767153](https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2024)767153)
- Drake, W. J. (2004). Reframing Internet governance discourse: Fifteen baseline propositions. In D. MacLean (Ed.), *Internet Governance: A Grand Collaboration* (pp. 122-161). United Nations Information and Communication Technology Taskforce.
- European Commission. (2022). *European Declaration on Digital Rights and Principles for the Digital Decade*. Publications Office of the European Union. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022DC0028>

- European Commission. (2022, December 14). *Commission proposes European Declaration on Digital Rights and Principles* (Press Release No. IP/22/452). European Commission - Press Corner. [https://ec.europa.eu/commission/presscorner/detail/en/ip\\_22\\_452](https://ec.europa.eu/commission/presscorner/detail/en/ip_22_452)
- European Commission & Council of the European Union. (2022, April 25). *Declaration for the Future of the Internet*. Publications Office of the European Union. <https://digital-strategy.ec.europa.eu/en/library/declaration-future-internet>
- European Union. (2012). *Charter of Fundamental Rights of the European Union*. Official Journal of the European Union, C 326, 391-407.
- Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *OJ L 119*, 4.5.2016, p. 1-88.
- 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). *OJ L 277*, 27.10.2022, p. 1-102.
- Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). *OJ L 2024/1689*, 12.7.2024.
- Santos, B. de S. (2020). *Toward a new legal common sense: Law, globalization, and emancipation* (3rd ed.). Cambridge University Press.
- United Nations. (2024, September 22). *Global Digital Compact*. Office of the Secretary General's Envoy on Technology. <https://www.un.org/digital-emerging-technologies/global-digital-compact>
- U.S. Department of Commerce. (1997). *A framework for global electronic commerce*.
- Working Group on Internet Governance. (2005). *Report of the Working Group on Internet Governance*. United Nations. <https://www.wgig.org/docs/WGIGREPORT.pdf>