



(RESEARCH ARTICLE)



## Regulating Platform Content Governance at Scale under the EU Digital Services Act (DSA) and the DSA Transparency Database

Jimmy Kinyonyi Bagonza \*

*Department of Information Technology, School of Computer Science and Information Sciences, University of the Cumberlands, Williamsburg, KY, USA.*

World Journal of Advanced Research and Reviews, 2026, 30(02), 2684-2695

Publication history: Received on 21 April 2026; revised on 27 May 2026; accepted on 30 May 2026

Article DOI: <https://doi.org/10.30574/wjarr.2026.30.2.1545>

### Abstract

The European Union Digital Services Act (DSA), set to be fully implemented in 2024, represents a significant transformation in the governance of digital platforms, marking a clear shift from voluntary, multistakeholder frameworks to a binding regulatory regime backed by financial penalties of up to 6% of global turnover. At the heart of this governance structure is the Statement of Reasons (SoR) Transparency Database, a publicly accessible, machine-readable repository that catalogs content moderation decisions made by Very Large Online Platforms (VLOPs) serving over 45 million EU users. This paper explores how the DSA's mandatory transparency framework shapes the strategic behavior of globally operating platforms, maps the evolving power dynamics among key stakeholders, including the European Commission, VLOPs such as TikTok, X, YouTube, and Meta, civil society organizations, academic researchers, and EU member states, and assesses the policy's broader global implications. Drawing on an empirical analysis of 439 million SoRs and a synthesis of contemporary legal scholarship, the study reveals that 99.13% of moderation decisions are applied uniformly across the EU/EEA, indicating a preference among platforms for industrial-scale efficiency over regionally tailored governance. The case of X (formerly Twitter) illustrates how mandatory transparency measures can bridge informational asymmetries by revealing systematic discrepancies in self-reporting. Furthermore, the analysis identifies macroeconomic repercussions for global technology supply chains, including a regressive compliance cost structure that reinforces the dominance of established players while disadvantaging challenger platforms from the Global South. The DSA's 'Brussels Effect' signals a global shift towards stringent legal frameworks in platform governance, with substantial implications for digital sovereignty, multilingual content moderation, free expression, and equitable participation in the global digital economy.

**Keywords:** Digital Services Act; Platform governance; content moderation; Transparency database; Brussels Effect; Very large Online Platforms; Platform Observability; Algorithmic accountability

### 1. Introduction

The governance of digital platforms has experienced a profound transformation over the last two decades. From around 2005 to 2020, European platform policy was marked by informal codes of conduct, industry self-regulation, and voluntary bilateral agreements between states and platform companies. As demonstrated by Gorwa (2019) [1] through the Platform Governance Triangle framework, these arrangements systematically concentrated regulatory authority within the state-firm dyad, effectively marginalizing civil society and leaving platforms with near-total informational asymmetry regarding their own moderation practices. Critics have consistently argued that this framework favored platforms' commercial interests at the expense of democratic accountability, user rights, and equitable access to speech.

\* Corresponding author: Jimmy Kinyonyi Bagonza.

The full implementation of the EU Digital Services Act (DSA) in 2024 represents a significant change in the regulatory landscape, comparable to the impact of the E-Commerce Directive of 2000. Unlike previous voluntary guidelines, the DSA imposes binding legal obligations on any platform that serves over 45 million users within the EU, regardless of the company's country of origin. This extraterritorial scope has forced US-based companies such as Meta, Google, Snap, and LinkedIn, as well as platforms linked to Chinese firms like TikTok, to fundamentally reshape their global content governance frameworks. A key component of the DSA is the Statement of Reasons (SoR) Transparency Database, which serves as an openly accessible, machine-readable repository of content moderation decisions. This database serves as the cornerstone of a new 'platform observability' framework, [2] facilitating near-real-time oversight by researchers, regulators, and civil society organizations.

This paper examines the Digital Services Act (DSA) as a case study in global IT policy through four key dimensions. First, it contextualizes the Act within the historical development of platform governance, highlighting the transition identified by Gorwa [1] from informal regulatory negotiations to formalized legislation. Second, it identifies the key actors, structural power dynamics, and governance tensions that the DSA framework has activated. Third, it evaluates the policy's empirical performance, utilizing a comprehensive analysis of the Statement of Reasons (SoR) database. Fourth, it considers the Act's broader global significance, including its macroeconomic impact on digital supply chains and its implications for platform governance in the Global South. This analysis primarily relies on the empirical study by Papaevangelou and Votta (2025) [2], which examines 439 million SoRs collected over a four-month period from September 2023 to January 2024, supplemented by the primary legislative text and relevant legal and computational scholarship.

The paper is structured as follows: Section 2 provides the historical and theoretical context. Section 3 analyzes the DSA's core governance innovation. Section 4 explores platform-specific compliance strategies. Section 5 maps the actor landscape and examines power dynamics. Section 6 discusses the role of civil society and academic research. Section 7 evaluates the global significance and macroeconomic implications. Section 8 addresses ethical considerations and the long-term effects of governance. Finally, Section 9 concludes with reflections on the future of platform governance regulation.

---

## 2. Theoretical Framework and Historical Context

### 2.1. Platform Governance Triangle

Gorwa's (2019) [1] Platform Governance Triangle serves as the essential theoretical framework for this analysis. By adapting Abbott and Snidal's (2009) [3] tripartite Governance Triangle, which classifies regulatory initiatives according to the relative influence of states, firms, and civil society/NGOs, Gorwa shows that EU content governance initiatives from 2005 to 2020 were predominantly concentrated within the state-firm dyad. Examples such as the EU Code of Conduct on Hate Speech (2017), the Code of Practice on Disinformation (2018), and earlier Notice and Action framework proposals exemplify this trend: bilateral negotiations between the European Commission and platform firms occurred largely without substantial participation from civil society, with enforcement relying more on reputational rather than legal mechanisms.

This institutional setup led to what Gorwa describes as 'legitimation politics.' Platforms participated in governance processes not to embrace genuine accountability but to influence oversight terms in ways that maximized their operational discretion. The voluntary compliance regimes allowed platforms to define their own success metrics, determine their audit methodologies, and selectively report outcomes. The Commission, lacking formal enforcement authority over content governance, found itself structurally dependent on platforms' cooperation, an asymmetry that systematically skewed policy outcomes in favor of platform interests.

The Digital Services Act (DSA) represents a significant state-led shift in the balance of the governance triangle since the E-Commerce Directive. Instead of entirely eliminating the informal governance layer, it supplements it with binding obligations, a dedicated enforcement mechanism, and a mandatory data infrastructure. The Facebook Oversight Board, the Global Network Initiative, and the Santa Clara Principles continue to operate alongside the DSA, creating a complex, overlapping governance landscape. However, the DSA's introduction of financial penalties of up to 6% of global annual turnover (DSA Article 52(3)) has fundamentally transformed the incentive structure governing how platforms engage with these voluntary frameworks.

## **2.2. From Transparency to observability: The conceptual advance**

A significant conceptual contribution of the DSA scholarship is the distinction between transparency and observability. Rieder and Hofmann (2020) [9] first articulated 'platform observability' as a governance goal that goes beyond mere disclosure. While transparency involves platforms releasing information about their practices, observability refers to the ability of external actors to monitor, interrogate, and verify platform behavior in near real time. Leerssen (2024) [4] traces this conceptual trajectory within the DSA, arguing that the SoR database marks the first legislative effort to operationalize observability at a regulatory scale.

This distinction is crucial because transparency without observability can and historically has been easily exploited for performative compliance. Kaushal et al. (2024) [5] demonstrate, through legal and empirical analysis, that the self-reporting nature of SoR submissions creates systemic accountability gaps. Platforms retain considerable discretion over how they classify decisions, which actions they categorize as 'content moderation,' and the level of detail included in their submissions. In the absence of external verification mechanisms, the database risks becoming an infrastructure for the appearance of accountability rather than for genuine accountability. The X case, discussed in Section 4.3, provides a concrete illustration of this risk.

---

## **3. The DSA's Core Governance Innovation: The Statement of Reason Database**

### **3.1. Architecture and Design Logic**

The Statement of Reasons database serves as the primary operational tool for platform observability under the DSA. According to Article 17, Very Large Online Platforms (VLOPs) are obligated to provide users with a statement of reasons whenever they restrict access to content. This statement must include details about the legal or contractual basis for the restriction, the geographic scope of its application, and whether automated tools were employed in the decision-making process. These statements are compiled into a publicly accessible, machine-readable database maintained by the European Commission and hosted at [transparencydatabase.eu](https://transparencydatabase.eu).

The design of this database embodies a deliberate strategy of delegation. By making moderation decisions machine-readable and openly accessible, the Commission has effectively outsourced the identification of compliance inconsistencies to the broader research community, a model described by Papaevangelou and Votta (2025) [2] as 'delegated surveillance.' This approach allows the Commission to supplement its investigative capabilities with insights from researchers and civil society, generating evidence relevant to enforcement. As a result, this model alleviates the Commission's direct data processing burden while significantly increasing the number of effective oversight participants.

However, Leerssen (2024) [4] identifies two notable architectural limitations within the current design. First, the database captures only a subset of moderation decisions, specifically those affecting particular pieces of content or user accounts, and does not mandate reporting on algorithmic curation decisions, adjustments to recommendation systems, or demotion practices, which may have similar or even greater impacts on content visibility. Second, the granularity of required disclosures varies across platforms, reflecting the discretion they exercise in interpreting their reporting obligations. Both limitations diminish the database's potential to support comprehensive accountability.

### **3.2. Empirical Performance: Evidence from 439 Million SoRs**

Papaevangelou and Votta's (2025) [2] analysis of 439 million Submissions of Reports (SoRs) submitted between September 2023 and January 2024 represents the most comprehensive empirical assessment of the database to date. Their findings reveal several noteworthy patterns that elucidate both the strengths and limitations of the observability framework.

Most prominently, the data indicate that 99.13% of all moderation decisions across the five major Very Large Online Platforms (VLOPs) were applied uniformly across the EU/EEA, devoid of geographic differentiation. This near-absolute uniformity suggests that these platforms operate industrial-scale content moderation systems that treat the entire EU as a single jurisdiction, despite considerable variation in national legal frameworks governing illegal content. The practical consequence is that moderation decisions are optimized for the largest or most legally stringent member states, primarily Germany and France, and are thus applied uniformly across jurisdictions with divergent legal thresholds and cultural contexts.

TikTok submitted the highest volume of SoRs, accounting for 235 million submissions and representing 53.54% of the total corpus. This statistic reflects the platform's reliance on automated detection systems that facilitate moderation

decisions on a scale unattainable through human review alone. Meta's submissions constituted the second largest volume, while X submitted only 466,000 SoRs, a figure later shown by Trujillo et al. (2025) [6] to be inconsistent with X's own transparency reports, which documented over 2 million moderation actions during the same period. YouTube's data revealed a distinct pattern: intellectual property violations accounted for a significant share of third-party flagged content, raising substantive questions about the influence of rights holders on the moderation architecture.

The study also identified a notable deviation from the uniformity pattern, with France emerging as an outlier for TikTok. In this jurisdiction, certain content categories exhibited enforcement delays exceeding 30 days, indicating that national regulatory pressures from the Commission Nationale de l'Informatique et des Libertés (CNIL) and French content regulators have engendered behavioral differentiation not evident in other jurisdictions. This finding is significant because it implies that the uniformity mandated by the Digital Services Act (DSA) is not technologically predetermined; rather, platforms can and do modify their moderation practices in response to credible national enforcement pressures. However, such differentiation requires sustained regulatory engagement that goes beyond the baseline requirements set by the Commission.

---

## **4. Platform Compliance Strategies and Strategic Behavior**

### **4.1. US-based Platforms: Governance and Bifurcation**

For US-based Very Large Online Platforms (VLOPs), compliance with the Digital Services Act (DSA) has resulted in a clear structural bifurcation of content governance systems. These platforms now operate with parallel regulatory frameworks: one tailored to meet the EU's mandatory System of Record (SoR) reporting obligations, the algorithmic disclosure requirements outlined in DSA Article 27, and the stricter moderation standards defined in Articles 16 to 20; the other tailored for their domestic US operations, which remain subject to the significantly more lenient Section 230 provisions of the Communications Decency Act. This bifurcation is not merely a matter of administration; it signifies a fundamental divergence in the regulatory philosophies underpinning the two systems. Section 230 views platforms as neutral conduits, granting them broad immunity for their editorial discretion, whereas the DSA categorizes them as systemic intermediaries with explicit responsibilities to manage risk and ensure procedural fairness.

Meta's public regulatory filings and DSA transparency reports illustrate the tangible operational impacts of this bifurcation: the establishment of dedicated EU Trust and Safety teams, the creation of EU-specific algorithmic disclosure systems, the redesign of appeals processes to comply with DSA Article 20, and the appointment of a legal representative in the EU pursuant to DSA Article 13. These structural changes incur high costs that do not translate proportionately into revenue within the EU, leading industry analysts to characterize this phenomenon as a 'regulatory overhead premium.' The competitive implications are substantial; larger incumbents can absorb this premium as a fixed cost of market entry, while smaller challengers face disproportionately greater burdens relative to their resource capabilities.

### **4.2. TikTok: Geopolitical Compliance and Industrial Scale Automation**

TikTok's compliance trajectory highlights the convergence of regulatory obligations and geopolitical strategy, setting it apart from its US-based counterparts. Faced with the dual pressures of Data Services Act (DSA) compliance mandates from the European Commission and ongoing national security concerns raised by EU member states and the US government regarding ByteDance's Chinese ownership, TikTok has initiated Project Clover. This European data localization initiative involves establishing dedicated data centers within the EU and creating a European Technology Security Governance Board to oversee data access. This initiative transforms DSA compliance infrastructure from a mere regulatory obligation into a geopolitical positioning tool, designed to both satisfy Commission requirements and demonstrate independence from Chinese state oversight.

As the largest single contributor of Statements of Reasons (SoRs), with 235 million submissions accounting for 53.54% of the total corpus [2], TikTok's moderation profile serves as a key benchmark for evaluating other platforms. Its uniformity rate exceeding 99% across the European Economic Area (EEA) results from highly automated detection systems that handle content at a scale that is largely incompatible with meaningful human review for most decisions. The strategic decision in 2024 to lay off approximately 300 human moderators at its Amsterdam facility and to redirect investment toward expanding automation infrastructure reflects a deliberate optimization strategy: minimizing compliance labor costs while ensuring the throughput necessary to meet SoR submission requirements. While this trade-off may appear rational from a cost-efficiency standpoint, it inherently diminishes the contextual sensitivity of moderation decisions, particularly for content in low-resource languages and culturally specific contexts, which automated systems struggle to manage effectively.

### **4.3. X (formerly Twitter): Adversarial Non-Compliance and the Limits of Self-Reporting**

X's approach to DSA compliance under its post-2022 ownership and management serves as a pivotal test case for the credibility of the observability framework's enforcement. Following the platform's restructuring, X adopted a distinctly adversarial stance toward Commission oversight, publicly framing DSA obligations as violations of editorial freedom and resisting audit requests under DSA Article 37. This stance reflected a strategic calculation: that the Commission's enforcement proceedings would be drawn out, that the reputational impact on the platform's core user base would be minimal, and that portraying non-compliance as a free speech issue would garner political support in U.S. regulatory discussions.

However, this calculation proved to be a misjudgment regarding compliance. In their 2025 study, Trujillo et al. [6] cross-referenced X's Statements of Reasons (SoRs) with the platform's transparency reports, uncovering a fundamental inconsistency: X's SoRs reported only manual moderation decisions, while its transparency reports acknowledged the extensive use of automated detection tools. Furthermore, X submitted only 466,000 SoRs during a period in which its own data indicated over 2 million moderation actions, a discrepancy of more than fourfold that cannot be attributed to definitional ambiguity. These inconsistencies, brought to light by independent researchers using the Commission's own database infrastructure, provided the evidence necessary for formal enforcement proceedings under DSA Article 66.

X's case has since become a paradigm in the DSA literature, illustrating how mandatory transparency mechanisms can counteract the informational asymmetry that platforms have historically exploited. By making self-reported data publicly accessible and machine-readable, the SoR database has empowered researchers to identify contradictions that the Commission's investigative resources could not have detected with the same speed or scale. The tactical implications are significant: platforms that underreport or misclassify SoR submissions face not only regulatory risks but also reputational risks stemming from independent research, creating a multi-layered accountability dynamic that enhances the Commission's enforcement capabilities.

### **4.4. YouTube (Google/ Alphabet): Hybrid Moderation and Rightsholder Asymmetry**

YouTube's compliance profile is notably different from TikTok's due to its hybrid moderation architecture, which integrates automated detection with manual review processes, resulting in more territorially differentiated outcomes. YouTube boasts the largest user base among designated Very Large Online Platforms (VLOPs) in the EU, with 416.6 million monthly active users. It faces particularly intricate moderation challenges given the diverse range of content categories present on its platform, including music, news, political commentary, and user-generated entertainment. In certain categories, the median response time for country-specific manual review decisions exceeds 100 days, reflecting both the high volume of requests and the complexity of applying jurisdiction-specific legal standards to ambiguous content.

The most noteworthy finding from YouTube's Statement of Responsibilities (SoR) data concerns the nature of third-party flagged content: intellectual property violations, primarily initiated by music labels, film studios, and other commercial rightsholders, account for the majority of third-party notifications. Papaevangelou and Votta (2025) [2] identify this as an emerging structural concern; the Digital Services Act's (DSA) procedural framework, designed to enhance accountability for harmful content, seems to have established a compliance channel that disproportionately benefits commercially organized rightsholder interests. Consequently, this creates an enforcement asymmetry in which well-resourced commercial entities can use the DSA's notification mechanisms more effectively than individual users or civil society groups seeking redress for content that affects public-interest concerns.

---

## **5. Actor Mapping and Power Dynamics**

Table 1 provides an in-depth analysis of the key stakeholders within the governance ecosystem of the Digital Services Act (DSA), systematically categorized by actor type, operational tier, and their principal governance functions.

**Table 1** Key Actor overview in the DSA Governance Ecosystem

<b>Actor</b>	<b>Type</b>	<b>Level</b>	<b>Primary Role</b>
European Commission (DG CNECT)	Regular / Executive	Supranational	Legislative Architect & enforcement authority
European Parliament	Legislator	Supranational	Co-legislator, democratic legitimization
EU Member States (Germany, France, Netherlands, Italy)	State Regulators	National	National digital coordinators content law
Meta- Facebook / Instagram	Platform firm	Transnational	High volume SoR submitter, Compliance actor
TikTok	Platform firm	Transnational	Dominant SoR Volume. Automated moderation
X formerly Twitter	Platform firm	Transnational	Self-reporting anomalies, enforcement target
YouTube – Google	Platform firms	Transnational	Hybrid moderation. IP dominated removals
Non – EU Governments	External Actors	International	Regulatory learning, Brussels effect recipients, Geopolitical Stakeholder
EFF, Access Now, Article 19	Digital Rights NGOs	Transnational	Advocacy, Watchdog, Litigation
Vetted Academic Researchers	Knowledge Actors	Transnational	Empirical accountability, DSA Art. 40 access
Content Right Holders	Industry interest	Transnational	IP enforcement influence on moderation
Human Moderators / Platform labor	Affected workers	Global	Operational Capacity. Labor rights concerns
End Users	Rights Holders	National / Global	SoR recipients. Redress mechanism beneficiaries

**5.1. Power Dynamics and Governance Phase Analysis**

Table 2 delineates the evolution of power dynamics across four distinct governance phases, highlighting the gradual curtailment of platform discretion alongside a concurrent increase in leverage held by both the state and the research community.

**Table 2** Governance Phase Analysis and Power Trajectories (2005 – Present)

<b>Phase</b>	<b>Coalition Dynamics</b>	<b>Power Trajectory</b>	<b>Enforcement Mechanism</b>
Pre-DSA (2005 – 2020)	State-Firm dyad dominates, informal codes of conduct, civil society largely excluded	VLOPs hold near absolute informational asymmetry, commission limited to soft persuasion	Voluntary compliance, no enforceable penalties, platforms set own transparency standards
DSA Negotiation (2020-2022)	Trilogue narrows VLOP discretion, civil society input via parliament, platforms lobby aggressively on SoR scope	Commission gains legislative initiative, VLOPs forced into formal regulatory engagement	Binding SoR obligation established, Art. 40 researcher access codified, Penalty framework agreed
Early enforcement (2022 -2024)	Commission designates VLOPs, National Digital Coordinators activated, X proceedings initiated,	State enforcement authority operationalized, platforms face	Platforms restructure global compliance pipelines, SoR

	<i>researcher outputs feed enforcement</i>	<i>financial penalty exposure for first time</i>	<i>database goes live, first formal proceeding opened.</i>
<i>Current Phase (2024 – present)</i>	<i>Empirical accountability ecosystem emerges; researcher findings inform enforcement actions, Civil Society litigation increases.</i>	<i>Commission uses third-party evidence for enforcement; the gap between performative and substantive compliance is now a central enforcement concern.</i>	<i>Automation is rapidly increasing among Very Large Online Platforms (VLOPs). The Article 40 vetted researcher program has been successfully implemented, and enforcement precedents have been established.</i>

## 5.2. Structural Tensions and Conflict Lines

The DSA governance framework creates several persistent structural tensions that significantly influence both compliance behavior and enforcement dynamics. The most fundamental of these tensions revolves around the classification of automation. Very Large Online Platforms (VLOPs) challenge the European Commission's interpretation of what constitutes an 'automated' moderation decision for the purposes of Significant Occurrence Reporting (SoR). This disagreement has considerable implications for compliance costs. Platforms contend that hybrid decisions, in which algorithmic signals prompt human review, should not require individual SoRs. However, the Commission's interpretive guidance indicates a broader reporting obligation. This definitional conflict directly impacts the comparability of data across platforms and undermines the reliability of the database as an accountability tool.

The second significant tension pertains to jurisdictional fragmentation. The Digital Services Act (DSA) harmonizes procedural obligations but explicitly assigns the definition of 'illegal content' to member state law (DSA Recital 12). As a result, platforms must navigate 27 distinct national legal frameworks governing content illegality, all within a unified procedural framework. This creates compliance complexities that disproportionately burden smaller platforms and incentivize larger Very Large Online Platforms (VLOPs), which already possess jurisdiction-specific legal teams, to preserve their structural advantage. Notable regulations such as Germany's NetzDG, France's Avia Law framework, and the Netherlands' platform liability provisions each impose requirements that interact with DSA obligations in ways that lack coherence, resulting in compliance uncertainty that benefits incumbents with the resources to manage this complexity.

The third tension involves access asymmetries within the accountability ecosystem. While Article 40's vetted researcher program is formally innovative, it has unintentionally established a two-tier knowledge infrastructure. Well-resourced academic institutions in Western Europe and the United States can effectively navigate the application processes, meet data protection requirements, and possess the technical capacity to manage large-scale datasets. In contrast, independent researchers, civil society organizations with limited technical capabilities, and institutions from the Global South encounter substantial barriers to access. These barriers mirror the same asymmetries the DSA aimed to address within the commercial sector. These dynamic risks concentrate the accountability function within a narrow epistemic community whose priorities may not align with the diverse needs of all affected communities.

## 6. Civil Society, Academic Research, and the Accountability Ecosystem

### 6.1. Civil Society Structural Position

Civil society organizations hold a position of structural ambiguity within the governance framework of the Digital Services Act (DSA). On the one hand, the DSA formally integrates civil society into its governance architecture in ways unprecedented in previous EU platform regulations. Provisions such as Article 40's research access, Article 14's trusted flaggers program, and the requirement under Article 34 for platforms to consult with civil society during risk assessments represent meaningful advancements in civil society participation. These provisions were largely achieved through parliamentary advocacy during trilogue negotiations, in which coalitions of NGOs, including Access Now, Article 19, and the Electronic Frontier Foundation, successfully lobbied for the inclusion of research access provisions that were absent from the Commission's original proposal.

Conversely, Gorwa (2019) [1] notes that civil society organizations often encounter a significant structural capacity gap between achieving advocacy success and exerting implementation influence. Gaining formal recognition within governance frameworks does not inherently provide the operational resources necessary to wield that influence effectively. This gap is particularly pronounced in the context of the Digital Services Act (DSA). The trusted flaggers

program mandates formal designation by National Digital Coordinators, a process that tends to favor established organizations with legal expertise and administrative resources over grassroots groups or community-based organizations. Similarly, the Article 40 research access program demands advanced data protection compliance and technical analytical skills. Digital rights organizations based in the Global South face additional challenges, including language barriers, a lack of legal recognition, and limited institutional access, which further hinder their ability to participate in a governance framework that is ostensibly designed to ensure accountability for global platforms.

## 6.2. Academic Research as Enforcement Infrastructure

Academic researchers have assumed an increasingly pivotal role within the governance ecosystem of the Digital Services Act (DSA), effectively serving as 'third-party enforcement infrastructure.' The model of delegated surveillance employed by the Commission, which renders platform data publicly accessible while relying on the research community to pinpoint inconsistencies, has fostered a structural dependency between the regulator and independent researchers, one without a clear precedent in prior regulatory frameworks. This dependency presents both advantages and potential vulnerabilities within the DSA's architecture.

The advantages are evident in specific cases, such as the investigation by Trujillo et al. (2025) [6], which cross-referenced SoR submissions with X's own transparency reports. This analysis produced compelling evidence of systematic underreporting that the Commission's audit capabilities would not have been able to replicate at a similar pace or scale. Additionally, the study by Papaevangelou and Votta (2025) [2], which analyzed 439 million SoRs, yielded empirical findings such as a 99.13% uniformity rate, delays in French enforcement, and the dominance of rightsholders in YouTube's data, which have directly influenced the Commission's enforcement priorities and sparked scholarly discussions regarding the sufficiency of the current design. Complementing this, Drolsbach and Pröllochs (2024) [10] provide further analysis of moderation patterns within the database, thereby enhancing the empirical accountability ecosystem.

However, the dependency also introduces structural vulnerabilities: the efficacy of DSA enforcement depends on the ongoing availability of research funding, the maintenance of researcher access under Article 40, and the ongoing engagement of academic institutions in data-intensive accountability efforts. Should funding priorities shift, platforms successfully challenge researcher access through legal means, or if the volume and complexity of SoR data surpass the analytical capacity of researchers, the generation of enforcement signals, upon which the Commission increasingly relies, could be substantially impaired. This structural dependency represents an underexamined risk within the DSA's design logic.

---

## 7. Global Significance and Microeconomic Implications

### 7.1. The Brussels Effect in Digital Governance

The global significance of the DSA is primarily manifested through what Bradford (2020) [13] describes as the 'Brussels Effect': a form of unilateral regulatory globalization that arises when the EU's market size, regulatory capacity, and credibility in rule-of-law converge to establish EU standards as the de facto global benchmark for firms operating within its jurisdiction. Helberger and Samuelson (2024) [7] further elaborate on this by suggesting that the DSA's Transparency Database could serve as the cornerstone of a global transparency regime in platform governance, representing the first regulatory instrument with both the legal authority and data infrastructure necessary to compel global platforms to systematically account for their moderation practices across a significant market.

The Brussels Effect operates through two complementary mechanisms in this context. The first is known as the compliance globalization effect: platforms that modify their moderation architectures to comply with the DSA find it operationally simpler to implement these systems globally rather than maintaining separate frameworks for EU and non-EU markets. Examples include Meta's EU-compliant appeals processes, TikTok's Project Clover data governance framework, and YouTube's DSA-specific algorithmic disclosure systems, investments that, once established, tend to be adapted for global operations. The second mechanism is referred to as the regulatory learning effect: non-EU jurisdictions actively draw inspiration from the DSA's design as a legislative template, tailoring it to their own constitutional and political contexts. Notable instances include the UK's Online Safety Act 2023, Australia's revised Online Safety Act, Canada's Bill C-63, and Brazil's PL 2630, each incorporating transparency and accountability mechanisms influenced by the DSA.

### 7.2. Geopolitical Tensions and Regulatory Fragmentation

Despite its potential to catalyze global regulatory convergence, the Digital Services Act (DSA) also contributes to regulatory fragmentation along geopolitical lines. The divergence between the EU and the US is particularly pronounced. The DSA's approach to content governance is deeply rooted in a European human rights tradition, grounded in the European Convention on Human Rights and the EU Charter of Fundamental Rights. This framework views freedom of expression as a qualified right that may be subject to proportionate restrictions in the interest of democratic values. In stark contrast, US First Amendment jurisprudence treats governmental interference with speech as presumptively unconstitutional and offers platforms near-absolute protection from liability for third-party content under Section 230.

This divergence has tangible implications for transatlantic coordination in platform governance. US platforms that comply with DSA requirements in the EU may confront potential conflicts with First Amendment arguments put forth by US legislators and advocacy groups, who perceive EU content governance obligations as an overreach beyond their borders. The ongoing regulatory scrutiny surrounding X's non-compliance has become a focal point in this broader geopolitical struggle, with some American commentators framing the enforcement actions by the Commission as an infringement on free speech rather than a legitimate exercise of regulatory authority. Furthermore, the TikTok case introduces an additional geopolitical layer: compliance with EU DSA requirements intersects with US national security concerns, complicating the platform's regulatory stance and exerting pressure for alignment choices that carry significant commercial implications.

### 7.3. Macroeconomic Impacts on Global Technology Supply Chains

Table 3 illustrates the macroeconomic ripple effects of the DSA across four interconnected domains of global technology supply chains, highlighting the differing impacts on EU-adjacent markets compared to those in the Global South.

**Table 3** DSA Macroeconomic Impacts across Global Technology Supply Chain Domains

<b>Supply Domain</b>	<b>Chain</b>	<b>EU Market Impact</b>	<b>Global South Impact</b>	<b>Emerging Structural Effect</b>
<i>Trust &amp; Safety labor Market</i>		<i>Rapid growth in specialized compliance roles in EU adjacent markets – Ireland, Netherlands</i>	<i>Outsourced human moderation in Global south increasingly restructured or automated away</i>	<i>Rising demand for multilingual policy engineers, new certification ecosystems emerging</i>
<i>Cloud and AI infrastructure Spending</i>		<i>VLOPs invest heavy in EU-hosted moderation infrastructure to satisfy DSA data requirements</i>	<i>Third party AI moderation vendors benefit from compliance driven procurement cycles</i>	<i>Emerging markets lack equivalent compliance infrastructure widening the digital divide</i>
<i>Startup &amp; SME market access</i>		<i>Higher per user compliance costs create structural barriers for challenger platforms</i>	<i>Incumbents leverage existing moderation infrastructure as a competitive moat</i>	<i>Global South startups face double burden; local regulations plus extraterritorial EU obligations if serving EU users</i>
<i>Content Supply chains</i>		<i>Commercial rightsholders, such as music labels and film studios, gain unfair enforcement power through third-party notification systems.</i>	<i>The creator economy is facing increased uncertainty regarding automated takedowns, with limited options for recourse against wrongful removals outside of EU jurisdictions.</i>	<i>Cross-border content licensing negotiations increasingly shaped by moderation risk exposure and jurisdiction specific takedown liability</i>

The compliance cost structure inherent in the Digital Services Act (DSA) functions as a regressive regulatory burden across the global platform economy. Per-user compliance costs associated with Source of Revenue (SoR) reporting, algorithmic disclosure, and risk assessment obligations are inversely proportional to the scale of the platforms: the largest Very Large Online Platforms (VLOPs) can distribute these costs among hundreds of millions of users, while emerging platforms, particularly those from the Global South seeking access to the EU market, bear disproportionately heavy burdens. This framework is not coincidental; it embodies the DSA's alignment with the operational models of the

platforms it aims to regulate, effectively establishing VLOP-scale compliance infrastructure as the baseline expectation for all market participants.

Moreover, the DSA's impact on global content supply chains extends beyond the platforms to include content creators, publishers, and rights-holding organizations. The prevalence of rightsholder notifications in YouTube's SoR data, a finding with significant implications for the creator economy, suggests that the DSA has unintentionally created an enforcement asymmetry that favors commercially organized content interests over individual creators and non-commercial expression. As cross-border content licensing negotiations increasingly hinge on moderation risk exposure rather than market-driven demand, the macroeconomic ramifications ripple across the entire global creator economy in ways that the DSA's architects did not fully foresee.

---

## **8. Ethical Dimensions, Privacy Concerns, and Long-Term Governance Effects**

### **8.1. The Ethics of Automated Governance**

The automation of content governance at an industrial scale raises significant ethical questions, which the DSA's observability framework highlights but does not resolve. The SoR database enhances accountability for individual moderation decisions while simultaneously aggregating data that may reveal sensitive patterns regarding user communities, speech behavior, and enforcement targeting. This openness of the database, an advantage from an accountability standpoint, also presents risks of re-identification and the unintended extraction of insights about specific communities or political groups, contrary to the intentions of the DSA's creators. Although GDPR compliance requirements address some of these issues, the interplay between the DSA's transparency obligations and GDPR's data minimization principles creates unresolved legal tensions that have yet to be adjudicated.

More fundamentally, the automation of content moderation reduces human expression to a classification problem suitable for machine learning, an approach that embeds specific assumptions about the nature of speech, the definition of harm, and the importance of cultural context. These assumptions are neither neutral nor universally applicable. Nicholas and Bhatia (2023) [8] document the consistent underperformance of large language models and automated detection systems for non-English content, particularly in low-resource languages, many of which are widely spoken in Africa, South Asia, and Southeast Asia. Moreover, the SoR database lacks a mandatory reporting requirement for the language of moderated content, leaving no systematic means to assess whether automated moderation systems disproportionately impact speakers of these languages, a major oversight in a regulatory framework that asserts global applicability.

### **8.2. The Structural Invisibility of Human Moderators**

The SoR framework creates a significant group of workers who are structurally overlooked: the human moderators whose labor is essential for content governance on platforms that employ hybrid moderation models. Many of these workers are engaged through outsourcing arrangements with subcontractors located in Kenya, the Philippines, and other regions in the Global South. They often face severe psychological harm due to repeated exposure to graphic, violent, sexual, and extremist content. Unfortunately, their working conditions, compensation structures, and psychological support resources fall outside the DSA's accountability framework, which focuses on the outcomes of moderation decisions rather than the circumstances under which those decisions are made.

This structural invisibility highlights a significant limitation of the observability paradigm: it prioritizes measuring platform behavior from the regulator's perspective while failing to adequately account for the human costs of content governance production. The DSA's emphasis on procedural transparency, documentation of moderation decisions and their underlying rationale, the labor conditions, power imbalances, and psychological burdens that affect human moderators' ability to exercise contextual judgment, which automated systems cannot replicate. Future governance frameworks that build upon the DSA's structure must explicitly address this gap if they **are** to achieve genuinely comprehensive accountability in platform content governance.

### **8.3. Long-Term Governance Trajectory**

Keller's (2024, cited in Papaevangelou and Votta [2]) identification of a 'factory-like' approach to content moderation raises significant concerns about the alignment of industrialized compliance with the democratic values the DSA ostensibly aims to uphold. Should DSA-style regulation expand globally through the Brussels Effect, this industrial moderation model, characterized by automated, high-throughput decisions that apply uniformly with limited contextual awareness, might emerge as the default standard for content governance worldwide. This development raises questions about whether such a model adequately addresses the needs of multilingual and culturally diverse communities.

Moreover, the trajectory outlined by Rieder and Hofmann (2020) [9], moving from informal bargaining to algorithmic transparency to platform observability, suggests a gradual intensification of regulatory controls. This evolution may compel platforms to choose between true accountability and maintaining operational scale. If the costs associated with industrial-scale moderation increase significantly due to regulatory compliance requirements, platforms might encounter structural incentives to either reduce the volume of user-generated content they host, concentrate their efforts in markets with lower regulatory capacity, or invest in technical systems that superficially meet reporting obligations while minimizing substantive compliance. Each of these potential responses would signify a governance failure that the DSA's designers sought to avert, indicating that the current framework may need continuous recalibration as enforcement experiences accumulate.

---

## 9. Conclusion

The European Union's Digital Services Act (DSA) signifies a pivotal intervention in the governance of platform content, marking the most consequential shift since the E-Commerce Directive of 2000. It is arguably the most ambitious initiative to date aimed at instituting democratic accountability within the governance of global digital infrastructure. By mandating the availability of machine-readable and publicly accessible records concerning content moderation decisions, the DSA has established a multi-tiered accountability framework that incorporates regulators, researchers, civil society, and users. This framework systematically dismantles the informational asymmetries that platforms have capitalized on over the past two decades.

An examination of empirical data from 439 million Statements of Reasons (SoRs) reveals both the demonstrable strengths and structural limitations inherent in this architecture. On the positive side, the database has laid bare systematic contradictions in X's self-reported compliance posture, validated the uniformity of moderation practices across the European Economic Area (EEA) as empirically verifiable, and thus subject to challenge, and fostered the development of a research-driven accountability ecosystem that enhances the European Commission's enforcement capabilities. Conversely, the observed 99.13% uniformity rate underscores that industrial-scale compliance is precipitating governance deficits such as a loss of contextual sensitivity, the homogenization of enforcement practices, and the automation of decisions that necessitate cultural understanding, the very issues the democratic rationale underpinning the DSA sought to mitigate.

Three significant implications for future governance frameworks emerge from this analysis. First, the delegated surveillance model, which relies on outsourcing inconsistency detection to the research community, necessitates ongoing investment in Article 40, encompassing researcher access, data literacy, and independent research funding to retain its efficacy as the volume of platform SoRs continues to grow. Second, the regressive compliance cost structure embedded within the DSA's architecture demands explicit remediation if the framework is to achieve its objective of fostering a competitive and pluralistic digital marketplace, rather than merely entrenching existing Very Large Online Platforms (VLOPs). Third, the structural invisibility of human moderators, coupled with the systematic inadequacies of automated systems, particularly concerning low-resource languages, represents ethical blind spots that a truly global governance framework must not overlook.

The global significance of the DSA transcends European boundaries. Through the Brussels Effect, it is influencing content governance systems worldwide and facilitating a transition from soft governance to hard law in jurisdictions as diverse as the United Kingdom, Australia, Canada, Brazil, and, in a nascent form, India and South Africa. Whether this global diffusion leads to genuine improvements in platform accountability or merely replicates the DSA's structural limitations on a global scale will critically depend on whether subsequent frameworks heed the empirical lessons derived from the current design's performance. The fundamental question for the next generation of platform governance, whether the industrialized moderation model aligns with democratic values at a global scale, remains unresolved, and its answer will significantly shape the informational landscape of democratic societies for decades to come.

---

## Compliance with ethical standards

### *Acknowledgments*

The author expresses gratitude to the School of Computer and Information Science at the University of the Cumberlands for their guidance and constructive feedback on earlier versions of this analysis. Additionally, the author acknowledges the foundational empirical contributions of Papaevangelou and Votta (2025), Trujillo et al. (2025), and Kaushal et al. (2024), whose published datasets and analyses form the basis of the empirical dimensions presented in this paper.

### *Disclosure of conflict of interest*

The author confirms that there are no financial or non-financial conflicts of interest related to the research, authorship, or publication of this article.

### *Statement of ethical approval*

This study is an analytical review of publicly accessible legislative texts, published academic literature, and transparency data from open-access platforms. It did not involve any human participants, personal data collection, or experimental methods, and therefore did not require institutional ethical approval.

---

## References

- [1] Gorwa, R. (2019). The platform governance triangle: conceptualising the informal regulation of online content. *Internet Policy Review*, 8(2).
- [2] Papaevangelou, C., & Votta, F. (2025). Trading nuance for scale? Platform observability and content governance under the DSA. *Internet Policy Review*, 14(3).
- [3] Abbott, K. W., & Snidal, D. (2009). The Governance Triangle: Regulatory Standards Institutions and the Shadow of the State. In *The Politics of Global Regulation* (pp.44-88).
- [4] Leerssen P. From Algorithmic Transparency to Platform Observability in the Digital Services Act. *Westminster Journal of Digital Studies*. 2024. doi:10.34669/wi.wjds/4.2.3
- [5] Kaushal, R., van de , K., Jacob, G., Catalina, S., & Gerasimos, I. A. (2024). Automated Transparency: A Legal and Empirical Analysis of the Digital Services Act Transparency Database. *FACCT '24: Proceedings of the 2024 ACM*.
- [6] Trujillo, A., Fagni, T., & Cresci, S. (2025). The DSA Transparency Database: Auditing Self-reported Moderation Actions by Social Media. *Proceedings of the ACM on Human-Computer Interaction*. *ACM Journals*.
- [7] Helberger, N., & Samuelson, P. (2024). The Digital Services Act as a Global Transparency Regime. *Verfassungsblog on Matters Constitutionasl*.
- [8] Nicholas, G., & Bhatia, A. (2023, May 23). Lost in Translation: Large Language Models in Non-English Content Analysis. *Centre For Democracy and Technology*.
- [9] Rieder B, Hofmann J. Towards platform observability. *Internet Policy Review*. 2020;9(4). doi:10.14763/2020.4.1535
- [10] Drolsbach, C. P., & Pröllochs, N. (2024). Content Moderation on Social Media in the EU: Insights From the DSA Transparency Database. *Association of Computing Machinery Digital Library*.
- [11] Husovec, M. (2024). Tamimng the Powers. *Principles of the Digital Services Act* pp. 3-18. Oxford University Press.
- [12] Klonick, K. (2017). The New Governors: The People, Rules and Processes governing Online Speech. *Havard Law School Review* 131(6) .
- [13] Bradford A. *The Brussels Effect: How the European Union Rules the World*. Oxford: Oxford University Press; 2020.
- [14] European Union. 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*. 2022 Oct 27;L277:1–102. Available from: <http://data.europa.eu/eli/reg/2022/2065/oj/eng>
- [15] United Kingdom. *Online Safety Act 2023*. Retrieved from *Online Safety Act 2023*: <https://www.legislation.gov.uk/ukpga/2023/50>