Journal of Law, Market & Innovation

Vol. 4 - Issue 2/2025

Alessandro Piovano[*]
Carlo Federico Vescovo[*]
Cristina Poncibò[*]

SPECIAL SECTION

AUTOMATING DSA ENFORCEMENT

A Socio-Technical framework for transparency compliance

Abstract

The European Union's Digital Services Act (DSA) is a landmark law aiming to make online platforms more transparent, accountable, and safe for users. But effective enforcement of the DSA poses significant challenges due to the scale of digital platforms and the complexity of their operations. This article presents a socio-technical legal framework designed to automate aspects of DSA enforcement, focusing on transparency obligations as a measurable and accessible starting point. The proposed framework combines legal analysis with computing techniques, such as web data extraction, natural language processing, and logic-based rule modelling, to continuously monitor platform compliance and designed to provide technological support to authorities and privates engaged in inspection and monitoring activities. By formalising DSA requirements into computable rules and developing tools to detect and report non-compliance, the approach seeks to bridge the gap between regulatory objectives and practical oversight capabilities. Case studies on selected DSA provisions (including obligations for contact points, terms of service clarity, transparency reports, notice-and-action systems, and advertising and recommender transparency) illustrate how the framework operates across different compliance areas. The article emphasises clarity and cross-disciplinary accessibility, aiming to foster dialogue between legal, policy, business, and technical stakeholders, and suggesting how regulatory automation tools can support authorities and platforms in upholding the DSA, and potentially other digital regulations, by providing scalable, objective, and transparent enforcement mechanisms.[1]

JEL CLASSIFICATION: K24, K23, C88, L86

SUMMARY

[*] PhD Student, Department of Law, University of Turin, Email: a.piovano@unito.it
[*] Research Assistant, Department of Law, University of Turin, Email: carlofederico.vescovo@unito.it
[*] Full Professor of Comparative Private Law, Department of Law, University of Turin, Email: cristina.poncibo@unito.it
[1] This work was developed as part of the PARASOL project at the Department of Law, University of Turin. For further information, contact the Authors.

DOI 10.13135/2785-7867/12271

1 Introduction

The Digital Services Act (DSA) represents a milestone in European digital regulation, introducing comprehensive rules to ensure that online platforms operate in a transparent, secure, and rights-respecting manner.[2] As part of a broader legislative package alongside the Digital Markets Act, the DSA's overarching goal is to create a safer and more accountable online environment while fostering innovation and competition in the EU digital market.[3]

The DSA emerged in response to some urgent issues identified over the past decade and a half, including the fragmentation of national regulations, extreme information asymmetries between users and platforms, challenges around meaningful digital consent, and the outsized influence of large technology companies on markets and society.[4]

Prior to the DSA, platforms operating across EU countries faced a patchwork of different rules, making compliance cumbersome and impairing smaller businesses.

The DSA, being an EU Regulation, directly harmonises these rules across Member States, aiming to reduce compliance burdens and provide uniform protections for consumers and businesses alike.[5] At its core, the DSA seeks to balance fundamental rights with technological innovation, indeed, EU legislators crafted the law to safeguard users' rights (such as freedom of expression and data protection) without unduly stifling the growth of digital services;[6] they also tackled pressing safety concerns, as the spread of illegal or harmful online content (like hate speech, disinformation, and counterfeit goods) had been exacerbated by social media and e-commerce growth.[7]

[2] Pietro Ghirlanda, 'How platform cooperatives can redress abuses of authority within digital markets' (2024) 3(3) Journal of Law, Market and Innovation 214.

[3] Andrej Savin, 'The EU Digital Services Act: Towards a More Responsible Internet' (2021) 04 CBS Law Research Paper 1; Andrea Turillazzi and others, 'The Digital Services Act: An Analysis of Its Ethical, Legal, and Social Implications' (2023) 15(1) Law, Innovation and Technology 83.

[4] Marsel Imamov and Natalia Semenikhina, 'The impact of the digital revolution on the global economy' (2021) 5(S4) Linguistics and Culture Review 968; Fischer-Lescano A and Teubner G, 'Regime-collisions: the vain search for legal unity in the fragmentation of global law' (2003) 25 Michigan Journal of International Law 999.

[5] Urbano Reviglio and Matteo Fabbri, 'The Regulation of Recommender Systems Under the DSA: A Transition from Default to Multiple and Dynamic Controls?' (*DSA Observatory,* 22 November 2024) <https://dsa-observatory.eu/2024/11/22/the-regulation-of-recommender-systems-under-the-dsa-a-transition-from-default-to-multiple-and-dynamic-controls/> accessed 20 June 2025.

[6] Savin (n 2).

[7] Nataliia Filatova-Bilous, Tetiana Tsuvina and Bohdan Karnaukh, 'Digital Platforms' Practices on Content Moderation: Substantive and Procedural Issues Proposed by DSA' in Conference on Integrated Computer Technologies in Mechanical Engineering – Synergetic Engineering (Springer Nature 2023).

To address this, the DSA imposes transparency obligations and "notice-and-action" mechanisms so that illegal content can be reported and removed quickly, with appropriate checks to protect lawful speech. Notably, the DSA positions the EU as a global leader in platform governance establishing standards for accountability and user empowerment that could influence internet regulation worldwide and the related economics trends[8]. The Regulation's implementation is thus entwined with Europe's "Digital Sovereignty"[9] ambitions to assert control over online harms and market fairness, and to export a values-based approach to digital governance.[10]

While the DSA's passage was a significant legislative achievement, enforcing its provisions effectively is an equally critical challenge yet to be resolved. Indeed, the law distinguishes obligations by platform size, subjecting Very Large Online Platforms (VLOPs, those with over 45 million EU users) to stricter requirements than smaller services, in line with proportionality principles,[11] acknowledging that the biggest platforms have the greatest impact and resources, and thus can bear more intensive compliance measures, whereas smaller businesses should not face undue burdens. Yet across all sizes, turning the DSA's legal mandates into practical reality will require new tools and strategies. Recognising this, our work proposes a technology-assisted enforcement framework, conceived as an auxiliary tool to facilitate the supervisory and verification functions of the competent authorities with respect to the provisions under examination, without in any way substituting their institutional prerogatives or discretionary decision-making powers.[12]

We build on the idea that "code can complement law"[13] by embedding regulatory checks and processes into digital systems; in other words, effective oversight of platforms may require automated or semi-automated systems working jointly with human regulators. This approach views the DSA not merely as a legal document, but as a system of rules that can be formally modelled and continuously monitored with the help of a software, designed as proposed in this article. By embracing innovations in data analysis and artificial intelligence for governance purposes, regulators can better keep pace with the fast-moving tactics of the online industry.

This article outlines a legal-informatics design for automated enforcement of key DSA transparency obligations, aiming to show the potential of the combination of legal research and computing techniques, which can make DSA compliance verification more systematic and scalable.

[8] Reviglio (n 4).

[9] Luciano Floridi, 'The fight for digital sovereignty: what it is, and why it matters, especially for the EU' (2020) 33 Philosophy & Technology 369.

[10] Turillazzi (n 2).

[11] Reviglio (n 4).

[12] Frank Pasquale, 'A rule of persons, not machines: the limits of legal automation' (2019) 87 (1) George Washington Law Review 1.

[13] Samer Hassan, Primavera De Filippi, 'The expansion of algorithmic governance: from code is law to law is code' (2017) 17 Field Actions Science Reports 88-90.

Following this introduction, Section 2 discusses the enforcement structure and the challenges that motivated an automated approach, highlighting issues such as algorithmic opacity and resource asymmetries. Section 3 presents our design, explaining why we focus on transparency-related duties and describing the technical components (from web data gathering to natural language processing and logical rule modelling) that make up the enforcement toolkit. In Section 4, we apply the framework on selected DSA provisions as case studies, illustrating how the framework has the potential to be the starting point to verify compliance with specific legal requirements (like providing proper contact points, publishing clear terms of service and transparency reports, handling user notices, and ensuring advertising and recommender system transparency). Section 5 concludes by reflecting on the benefits and limitations of this approach and its broader implications for digital governance. Throughout this work, we emphasise clarity and accessibility, aiming to inform a wide audience, including legal scholars, policymakers, technologists, and industry practitioners about how automated tools can support the DSA's successful implementation.

2 DSA enforcement structure and challenges

The Digital Services Act (DSA), formally Regulation (EU) 2022/2065, establishes a legal framework for the oversight and enforcement of digital services within the European Union.[14] This framework is characterised by a multilayered governance structure, allocating responsibilities between national authorities and EU institutions, with the objective of ensuring effective and coherent supervision of online intermediaries.[15]

At the national level, each Member State is required to designate a Digital Services Coordinator (DSC),[16] who acts as the competent authority responsible for monitoring compliance by service providers established within its territory. The DSCs are entrusted with investigatory powers, the ability to impose administrative sanctions, and the obligation to cooperate with other national coordinators and EU bodies to facilitate cross-border enforcement.[17]

At the supranational level, the European Commission retains exclusive supervisory and enforcement competences over Very Large Online Platforms (VLOPs) and Very Large Online Search Engines (VLOSEs), as defined by the Regulation. These entities, identified based on a minimum threshold of 45 million average monthly active users in the EU, are

[14] cf Recital 4 DSA.

[15] Jens-Peter Schneider, Kester Siegrist and Simon Oles, Collaborative Governance of the EU Digital Single Market established by the Digital Services Act (2023) 9 University of Luxembourg Law Research Paper 1.

[16] Un sito ufficiale dell'Unione europea <https://digital-strategy.ec.europa.eu/it/policies/dsa-dscs> accessed 20 June 2025.

[17] Petros Terzis, Michael Veale and Noelle Gaumann, 'Law and the emerging political economy of algorithmic audits' in Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency ((FACCT '24), June 03-06, 2024, Rio de Janeiro, Brazil) 1255.

subject to enhanced obligations due to their systemic relevance for the information environment, the digital economy, and the protection of fundamental rights.[18]

To ensure effective compliance with such obligations, the Commission is vested with a wide array of investigative and coercive powers. It may initiate formal proceedings against providers suspected of breaching the Regulation, upon notification to national coordinators and the European Board.[19]

Upon initiation of such proceedings, the Commission assumes a leading role and may temporarily suspend the supervisory competences of national authorities; indeed, the Commission may also request cooperation from Member State authorities in accessing documents, information, and premises located in their jurisdiction, insofar as they are relevant to the investigation.[20]

To collect evidence, the Commission may issue simple requests or binding decisions requiring the disclosure of information, these may be addressed to platforms and to third parties reasonably presumed to possess relevant data.[21]

Any such request must specify the legal ground and the purpose of the inquiry, the type of data required, the deadline for submission, and the consequences for failure to comply, which include financial penalties and daily fines, indeed, addressees are under a legal obligation to provide complete and accurate responses, and remain fully liable in case of omissions, delays or inaccuracies.[22]

The Commission may also demand access to platform databases and algorithms, as well as detailed technical explanations of their functioning. This kind of investigative actions may include the appointment of independent auditors and external experts, potentially in coordination with national authorities, who support the Commission in verifying compliance and ensuring impartial assessments.[23] Moreover, the Commission may require the preservation of technical documentation deemed necessary to assess regulatory implementation.

Where a breach of substantive or procedural obligations is established, the Commission may adopt a formal decision of non-compliance and order the provider to implement corrective measures within a specified period.[24] If the provider fails to comply, the Commission is empowered to impose financial penalties of up to 6% of the provider's total worldwide annual turnover, additional fines of up to 1% may be levied in instances of obstruction, misleading information, or non-cooperation during investigations.[25] This dual-

[18] cf Article 33 DSA.

[19] cf Article 66 DSA.

[20] Ilaria Buri and Joris van Hoboken, The DSA supervision and enforcement architecture (DSA Observatory 2022) 24.

[21] cf Article 67 DSA.

[22] Folkert Wilman, 'The Digital Services Act (DSA) - An Overview' [2022] SSNR https://ssrn.com/abstract=4304586 or http://dx.doi.org/10.2139/ssrn.4304586 accessed 26 June 2025.

[23] cf Article 72 DSA.

[24] cf Article 73 DSA.

[25] cf Article 74 DSA.

tier sanctioning regime reinforces the Commission's executive function and constitutes one of the most stringent enforcement mechanisms under EU digital regulation.[26]

To ensure institutional coordination, the DSA also establishes the European Board for Digital Services,[27] composed of the national DSCs and chaired by the Commission. The Board's tasks are various and include, especially, promoting the consistent application of the Regulation, exchanging best practices and tools, and issuing non-binding opinions on emerging regulatory challenges.[28]

Finally, the Regulation introduces a series of procedural safeguards aimed at reinforcing operational enforcement capacity. VLOPs and VLOSEs are required to grant access to data essential for compliance monitoring,[29] to undergo independent audits (cf Article 42 DSA), and to cooperate proactively with authorities to facilitate regular, transparent, and proportionate oversight.

This complex architecture reflects the DSA's ambition to address the challenges of the digital environment by combining decentralised supervision with centralised enforcement. However, as the subsequent analysis will demonstrate, the actual effectiveness of this framework depends heavily on the availability of adequate resources, specialised technical expertise, and innovative tools capable of responding to the dynamics of a rapidly evolving digital ecosystem.

In addition to public enforcement, the DSA assigns an active role to platforms themselves, which are required to implement mechanisms for content moderation and self-regulation these obligations constitute a form of private enforcement, whereby service providers must develop and manage tools for receiving notices of illegal content, act promptly, report the decisions taken, and ensure algorithmic transparency. This kind of co-responsibility implies that technological enforcement solutions must interact not only with public authorities but also with the internal systems operated by platforms.[30]

From the moment it was approved, a clear tension emerged between the DSA's regulatory goals and the practical realities of enforcement, indeed, several factors make traditional enforcement methods (eg, manual audits or complaint-driven investigations) often inadequate in the digital context[31]. First, modern platforms rely heavily on complex machine learning algorithms and deep neural networks to manage vast volumes of user content, giving rise to what scholars call "algorithmic opacity".[32] Furthermore, the

[26] Buri and Van Hoboken (n 19).

[27] Official site of European Board for Digital Services <https://digital-strategy.ec.europa.eu/en/policies/dsa-board> accessed 26 June 2025.

[28] cf Article 61 DSA.

[29] cf Article 40 DSA.

[30] Miguel Del Moral Sanchez, 'The devil is in the procedure: private enforcement in the DMA and the DSA' (2024) 9 University of Bologna Law Review 7.

[31] Afzal Jamil, "Digital Law Enforcement Challenges and Improvement" in *Implementation of Digital Law as a Legal Tool in the Current Digital Era* (Singapore: Springer Nature 2024) 47, 48.

[32] Motahhare Eslami, Kristen Vaccaro, Min Kyung Lee, Amit Elazari Bar On, Eric Gilbert, Karrie Karahalios, "User attitudes towards algorithmic opacity and transparency in online reviewing platforms" in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (New York, ACM 2019) 1, 14.

decision-making processes are often so complex, or kept proprietary, in a way that outsiders (including regulators) cannot easily understand or scrutinise them and in some cases, platforms intentionally design systems to be opaque or resistant to scrutiny. For example, through techniques of algorithmic laundering, a company might obfuscate how its content moderation AI works by continuously altering training data or model parameters, thereby thwarting external audits.[33]

The result is a large "grey area" where detecting violations of the DSA becomes exceedingly difficult without specialised tools. Another challenge is the global and distributed nature of online services: major platforms operate data centres and content delivery networks across multiple jurisdictions, which can facilitate regulatory arbitrage by letting companies locate certain business functions in countries with more lenient rules, impeding the enforcement of the DSA. This jurisdictional fragmentation allows companies to partially evade oversight, as noted by observers of EU digital regulation.[34] Additionally, the sheer disparity in resources between large tech companies and regulatory agencies raises concerns as Big Tech firms may employ hundreds of engineers and lawyers focused on content policies, whereas national regulators might have only a handful of technical experts at their disposal. This asymmetry means platforms can often adapt or reinterpret rules faster than authorities can monitor or respond. As a European Commission report pointed out, even well-intentioned regulations can fall short if enforcement bodies lack the technical tools and staff to keep up.[35]

The DSA's success hinges on addressing enforcement gaps created by modern technology, indeed, primary challenges include the opacity of algorithms and decision-making on platforms, which frustrates accountability, platforms' ability to exploit cross-border differences and technicalities to dodge compliance, and the limited capacity of regulators to perform large-scale, real-time supervision of platform activities. Recognising these challenges suggests that traditional enforcement must be augmented with automated, tech-assisted solutions.[36] If regulators can leverage advanced tools to inspect platforms continuously and objectively, they stand a better chance of ensuring that the DSA's provisions (for example, requirements about content moderation transparency or data access for researchers) are met in practice. These issues underscore why we view enforcement as the "Achilles' heel" of the DSA's otherwise robust legal framework, leading us to develop a methodology that directly tackles the enforcement challenge by combining legal criteria with computational monitoring.

[33] Meghan J Ryan, 'Secret algorithms, IP rights, and the public interest' (2020)21(1) Nevada Law Journal 61, 90.

[34] Caroline Cauffman, Catalina Goanta, 'A new order: The Digital Services Act and consumer protection' (2021) 12(4) European Journal of Risk Regulation 758, 774.

[35] Jamil (n 30).

[36] Suzanne Vergnolle, 'Enforcement of the DSA and the DMA – What did we learn from the GDPR?' in Heiko Richter, Marlene Straub, and Erik Tuchtfeld (eds), *To Break Up or Regulate Big Tech? Avenues to Constrain Private Power in the DSA/DMA Package* (Munich 2021) 103.

Building upon these challenges, it becomes evident that the architecture of the DSA, while ambitious, leaves open critical enforcement vulnerabilities that require innovative regulatory thinking.[37] Indeed, the reliance on procedural guarantees and transparency obligations may create a lack of compliance, where platforms formally adhere to reporting duties without substantially altering harmful practices, thus enabling platforms to meet the letter of the law while circumventing its spirit.[38]

In this regard, mere disclosure obligations such as the requirement to publish transparency reports or content moderation policies are insufficient if regulators lack the technical capacity to audit, verify, and interpret such disclosures effectively. Without tech-tools capable of parsing vast datasets, identifying inconsistencies, and cross-referencing publicly disclosed information against actual platform behaviours, the DSA's transparency measures risk becoming performative rather than transformative. Moreover, the asymmetry of information between regulators and platforms is compounded by the dynamic and evolving nature of algorithmic decision-making, unlike static compliance parameters in traditional industries, the digital ecosystem is characterised by constant iteration. Machine learning models undergo continuous retraining, and new recommendation strategies are deployed frequently, often without prior notice or public scrutiny; this algorithmic drift undermines the stability of compliance assessments, rendering periodic human audits obsolete almost immediately after completion.[39]

To further complicate enforcement, the DSA's regulatory framework faces the inherent challenge of legal ambiguity in defining key concepts. Terms such as "systemic risks", "appropriate content moderation", and "effective transparency" are open to interpretive variance, both in judicial application and in technical implementation. While this flexibility allows the DSA to remain adaptive to future technological developments, it also creates space for platforms to strategically interpret these obligations in ways that minimise compliance costs without necessarily advancing the Regulation's fundamental objectives.[40] This ambiguity has a direct impact on the enforceability of substantive rights under the DSA, indeed, without a precise normative framework that translates high-level legal concepts into measurable, operational criteria, the effectiveness of any enforcement, manual or automated, can be compromised. Additionally, the DSA does not fully account for the phenomenon of "compliance theatre" wherein platforms present curated datasets and controlled access to regulators and researchers, effectively shaping the narrative around their compliance efforts.[41]

[37] Ghirlanda (n 1).

[38] Florence G'sell, 'The digital services act (DSA): a general assessment' in Antje von Ungern-Sternberg (ed), *Content Regulation in the European Union - The Digital Services Act, TRIER STUDIES ON DIGITAL LAW*, Vol 1 (IRDT 2023).

[39] Christoph Busch, 'From algorithmic transparency to algorithmic choice: European perspectives on recommender systems and platform regulation' in S Genovesi, K Kaesling and S Robbins (eds), *Recommender Systems: Legal and Ethical Issues, The International Library of Ethics, Law and Technology*, Vol 40 (Springer 2023).

[40] U Kohl, 'Toxic recommender algorithms: immunities, liabilities and the regulated self-regulation of the Digital Services Act and the Online Safety Act' (2024) 16(2) Journal of Media Law 301.

[41] G'sell (n 37).

Considering these obstacles, enforcement bodies face a triple challenge: (i) legal ambiguity that complicates the translation of regulatory objectives into enforcement actions; (ii) informational asymmetries that impede the discovery of non-compliance; and (iii) resource constraints that limit their ability to keep pace with fast-moving technological changes.[42]

The solution, therefore, cannot rely solely on traditional supervisory practices or the goodwill of regulated entities, instead, it necessitates the integration of automated, technology-assisted enforcement mechanisms capable of continuous, objective, and scalable monitoring, indeed by embedding regulatory logic directly into computational systems, enforcement agencies can proactively detect patterns of non-compliance, assess the authenticity of transparency disclosures, and identify latent systemic risks without depending exclusively on platform cooperation.[43]

This approach anticipates the subsequent sections of our work, where we outline a design framework that formalises DSA obligations into machine-readable rules and employs advanced data extraction and natural language processing techniques to verify compliance autonomously. Specifically, by focusing initially on the DSA's transparency obligations, arguably the most objectively verifiable and publicly accessible set of rules, we establish a foundation for a broader framework capable of extending into more complex compliance areas, including content moderation practices, advertising transparency, and systemic risk mitigation. While the DSA represents a significant regulatory advancement, its enforcement success hinges on the development of technological infrastructures that complement legal mandates with real-time, automated oversight. Only through this synergy between law and technology can the European Union hope to close the enforcement gap and ensure that the DSA achieves its intended effect of creating a safer, more transparent, and accountable digital environment.

3 A design for automated DSA enforcement

3.1 The starting point: transparency obligations

As a starting point in automating the enforcement of the DSA, we decided to delimit the area of concern by focusing first on transparency obligations, the duties of platforms to disclose certain information about their operations to regulators and the public. We identified transparency-related requirements as an ideal starting point for automation design because they are among the most concrete and observable rules in the DSA. Unlike some obligations that might require subjective judgments or internal data (eg, assessing whether content moderation decisions were "appropriate"), transparency measures often

[42] ibid.

[43] Robert Mor and Johannes Dimyadi, 'The promise of automated compliance checking' (2021) 5 Developments in the Built Environment 1.

manifest as information that platforms must publish openly. This means compliance (or non-compliance) with these rules can be checked from an external perspective, including by users or automated systems, without special access to a company's internal databases, in other words, transparency provisions create data that is intentionally public-facing, which we can leverage for independent verification and to carry out preliminary tests aimed at assessing the effectiveness of the applied technological solutions. Transparency in the DSA is not just an abstract principle, but it is implemented through specific mandates, notably, platforms are required to publicly disclose key information about how they moderate content, how their recommender algorithms work, and how online advertising on their service is targeted and presented.[44]

For example, if a social media platform removes a user's post, the DSA obliges the platform to provide an explanation to the user, including the reason and the basis in their terms of service.[45] Likewise, large platforms must maintain advertising archives where details of ads (such as who paid for them and what targeting criteria were used) are available for scrutiny by anyone.[46] These transparency reports and databases are valuable because they offer observable indicators of compliance. By examining them, one can infer whether a platform is following the rules, for example by assessing whether the required information is provided in a clear manner and whether the content moderation reports are updated as mandated.

Focusing our design methodology on transparency obligations offered several advantages. First, as noted, verifying transparency does not require privileged access to a platform's back-end systems or personal user data, preserving user privacy. This increases the feasibility of independent oversight.[47] Second, transparency criteria are often binary or clearly defined whether an item (like a contact email or a summary of terms) is published or it is not; either a report contains certain statistics, or it does not, this allows us for objective checks. Third, many transparency duties apply across all platforms (with additional intensity for VLOPs but still relevant to smaller ones), meaning an automation approach here can scale to various contexts.[48] Finally, transparency requirements typically involve periodic disclosures (eg, monthly content moderation reports, continuously updated ad repositories), this creates a need for continuous monitoring, which is well-suited to automation. An automated system can be scheduled to regularly crawl and analyse the latest disclosures from platforms, catching compliance lapses (such as a report not being updated on time) much faster than occasional human audits could.

[44] Wilman (n 21).
[45] cf Article 14 DSA.
[46] Magdalena Knapp, Anna Piszcz, "Moving towards more transparent online platforms under the Digital Services Act" in Dušan V Popović and Rainer Kulms (eds), *Repositioning Platforms in Digital Market Law* (2024) 105, 123.
[47] Cauffman, Goanta (n 33).
[48] Amanda Reid, Evan Ringel, 'Digital intermediaries and transparency reports as strategic communications' (2025) 41(1) The Information Society 1, 18.

In our design, we thus narrowed the scope initially to a set of DSA provisions that revolve around transparency and accessibility of information; by doing so, we established a solid foundation of public data and clear-cut criteria on which to build automated enforcement tools. This choice is not meant to diminish the importance of other DSA facets (like risk assessments or crisis response duties), but rather to phase the development and proving the concept on transparency can pave the way to extending automated checks to more complex obligations in the future.[49] The next subsection details how we identified the specific articles to target and how those choices guided the technical implementation.

3.2 Scope and article selection

Within the DSA's many provisions, we selected a subset of articles that are both central to the Act's transparency goals and amenable to automated monitoring; this selection was guided by a dual logic: prioritising rules that have high regulatory importance and those that can be translated into clear computational checks. On the one hand, transparency is a cornerstone of the DSA's approach to balancing innovation with fundamental rights protection,[50] so it made sense to focus on articles enforcing transparency, on the other hand, each legal obligation is worded differently, some are straightforward (eg, "provide a point of contact"), while others are more qualitative (eg, "terms of service must be clear and understandable"). The first step was to ensure we target the right entities; Article 3 of the DSA defines which online services fall under which category (eg, intermediary services, hosting services, online platforms, VLOPs).[51] Any automated enforcement tool must incorporate this scope determination and checking whether a given website or service is subject to certain obligations, for instance, an obligation might apply only to "online platforms" but not to mere conduits (like ISPs). We included this classificatory step as a prerequisite in our framework, if a platform does not meet the DSA's definitions, our system is designed to recognise that and avoid a false non-compliance flag.

Next, we identified specific transparency obligations to analyse. Article 11 and Article 12 were chosen as a combined case focusing on points of contact, the first requires platforms to designate a single point of contact for communicating with regulators, and the second requires an electronic contact point for users, which must be easy to access and not purely automated (ie, users should be able to reach a human). These provisions are fundamental because if regulators or users cannot effectively contact a platform, enforcement of other rules becomes difficult, from an automation perspective, verifying compliance with Articles 11-12 is feasible by scanning the platform's website for contact info and testing its accessibility, for achieving this goal, we broke down this verification

[49] Knapp, Piszcz (n 45).
[50] Turillazzi (n 2).
[51] cf Article 3 DSA.

into specific metrics: (a) the contact information must be clearly visible and reachable within a couple of clicks from the homepage, (b) at least one non-automated channel (eg, a human-monitored email address or phone number) must be provided,[52] and (c) any descriptions around the contact must not be misleading or overly technical (to satisfy the "easily accessible" spirit). For example, the tool is designed to check that a user can navigate to the "Contact" or "Legal" page and find an email address without encountering login walls or obscure menus. We also included text analysis to ensure the language describing the contact is straightforward (no confusing jargon that might deter users). These criteria have the potential to translate the ambiguous terms "easily accessible" and "not exclusively automated" into concrete checkpoints that an algorithm can evaluate.[53]

We also targeted Article 14, which deals with terms and conditions transparency, under this article platforms must state their content moderation policies clearly in their terms of service, notify users of any significant changes to those terms, and ensure terms are appropriate for minors if the service is likely to be accessed by them. This was included because terms of service are a primary way platforms communicate rules to users, and lack of clarity here undermines user rights.[54] We identified multiple aspects of Article 14 to examine including whether the terms are presented in plain language, whether changes to the terms are announced or highlighted (as required), whether there is a concise summary of key points (the DSA encourages summaries for accessibility), and if the platform is known to be used by minors, whether the terms account for that (eg, simpler language or special sections.[55] These can be checked by analysing the text of the terms; in fact readability metrics can signal if the language is too complex, and a comparison of different versions of the terms over time (tracked via our tool) can show if changes were disclosed.[56]

Another important metric is whether a change log or notice is provided when the terms update, which we can detect by looking for dates or "last updated" notices and comparing content snapshots. An additional crucial area is transparency reporting, covered by Article 15, under this provision, larger online intermediaries must regularly publish reports with statistics on content moderation (eg, number of removal orders from authorities, number of user complaints, outcomes, etc). We chose Article 15 to see how our method can handle quantitative data and cross-referencing, for instance, if a platform's transparency report claims it handled a certain number of illegal content notices, our system could cross-check

[52] cf Article 12's requirement.

[53] Orlando Amaral Cejas, Muhammad Ilyas Azeem, Sallam Abualhaija and Lionel C Briand, 'NLP-based automated compliance checking of data processing agreements against GDPR' (2023) 49(9) IEEE Transactions on Software Engineering 4282.

[54] Katarzyna Wiśniewska and Przemysław Pałka, 'The impact of the Digital Content Directive on online platforms' terms of service' (2023) 42 Yearbook of European Law 388.

[55] cf Article 14 DSA.

[56] Marco Lippi, Przemysław Pałka, Giuseppe Contissa, Francesca Lagioia, Hans-Wolfgang Micklitz, Giovanni Sartor, Paolo Torroni, "CLAUDETTE: an automated detector of potentially unfair clauses in online terms of service" (2019) 27(32) Artificial Intelligence and Law 117, 139; Cauffman, Goanta (n 33).

consistency by verifying that each required category (eg, notices from governments vs notices from users) is present and that the report is updated on schedule.[57] We also consider the quality and granularity of the information, checking if each report breaks down the data as the DSA requires and if there are obvious omissions. Our design framework employs text parsing and pattern matching to locate relevant sections of the report and ensure key terms and figures are included. One of our future works is to understand how to cross-validate some figures with external data (for example, if the EU Commission publishes how many orders it sent to a platform, the platform's report should not conflict with that number).

We later considered Article 16, the "notice-and-action" mechanism, which requires platforms to provide easy ways for users to notify them of illegal content, and to act on those notices promptly while informing users of the outcome.[58] This provision is about the user experience of reporting content so, to automate enforcement, we examine whether the platform's interface offers clear and accessible reporting channels.[59] The system will simulate a typical user experience to assess the availability and accessibility of reporting mechanisms, for instance, the presence of a "Report" button or a dedicated form for notifying illegal content. It further examines compliance with procedural requirements, such as whether the platform issues confirmations or follow-up communications in response to user notices, as mandated by the regulation. Some of this can be inferred by analysing the help pages or terms (which should describe the notice process). Additionally, we check if multiple channels for reporting exist (webform, email, etc), since accessibility is improved by offering alternatives. While fully testing the responsiveness (eg, measuring actual removal times) is beyond a static crawl, our methodology flags whether the necessary systems appear to be in place and documented.

Finally, we addressed Articles 26 and 27, which focus on online advertising and recommendation algorithms transparency for large platforms, article 26 mandates that users be clearly informed when content they see constitutes advertisement and be able to access details such as who paid for the ad and why it was shown to them (targeting criteria). Article 27 requires platforms to disclose the "main parameters" of their content recommendation algorithms (for instance, the criteria that determine news feed rankings) and to offer users options to modify or opt out of personalised recommendations. These obligations are at the frontier of transparency, aiming to unveil algorithmic influence that users historically had little insight into; they are also among the hardest to monitor externally.[60] We approached them by focusing on whether the platform provides the required disclosures in an intelligible way. For ads (Art 26), our scraper checks if ads on a

[57] Reid, Ringel (n 47).
[58] cf Article 16 DSA.
[59] Daniel Holznagel, 'How to apply the notice and action requirements under Art. 16(6) DSA—Which action actually?' (2024) 25(6) Computer Law Review International 172,179.
[60] Vergnolle (n 35).

platform are labelled (like with “Sponsored” tags) and if clicking those labels leads to further information (many platforms have an “Why am I seeing this ad?” feature which should contain the mandated info).[61] We also look for the existence of the platform’s ad repository (for VLOPs, the DSA requires a public archive of all ads served). For recommender systems (Art 27), we search the platform’s user settings and help pages for explanations about how recommendations are generated and instructions for users to change their feed settings. A challenge here is that compliance might be formal rather than effective; a platform could nominally state some generic info about its algorithm without truly empowering users. To tackle this, our methodology incorporates an “explainability” check. For instance, we scan for whether concrete parameters (like “based on posts you liked” or “chronological feed option”) are mentioned as opposed to vague language.[62] We also note if the user interface allows switching off personalisation, as Article 27 effectively demands offering a non-personalised alternative.

By deliberately selecting Articles 11, 12, 14, 15, 16, 26, and 27, we created a testbed of varied transparency obligations, each bringing out different dimensions related to contact info availability, clarity of legal terms, quantitative reporting, interactive user-facing processes, and algorithmic transparency. The next step was to devise a unified method to automate the checking of all these elements. We proceeded to break down the enforcement verification into phases and components that could handle this diversity systematically.

3.3 Technical components of the framework

To implement the above enforcement checks, we developed a multi-phase design methodology integrating several technical instruments, in broad terms, the process involves: (i) gathering the relevant information from platforms (data collection), (ii) standardising and encoding the legal rules in a form a computer can work with, (iii) automatically analysing the collected data against those rules, and (iv) storing results and presenting them in a meaningful way, in the following sections, we describe each of these components and the technologies used.

3.4 Information gathering

In the absence of direct data-sharing interfaces from platforms, our system relies on web scraping to collect publicly available information.[63] Web scraping is an automated method of retrieving web pages and extracting specific content from them, for this we designed scrapers to target the sections of a platform’s website likely to contain

[61] Knapp, Piszcz (n 45).
[62] Hassan, De Filippi (n 12).
[63] Moaiad Ahmad Khder, ‘Web scraping or web crawling: State of the art, techniques, approaches and application’ (2021) 13(3) International Journal of Advances in Soft Computing & Its Applications 1, 12.

compliance-related information, for instance, the “Terms of Service” page, the “Transparency Report” page, the footer where contact details are often listed, and any dedicated “DSA compliance” or legal resources if provided by the platform. Using a combination of HTTP requests and browser automation, our tool aims to handle both static pages and those that require running JavaScript (some transparency portals are interactive or only load data via scripts). For static pages (like a simple terms of service text), an HTTP fetch followed by parsing the HTML for relevant sections (using selectors for headings like “Contact” or keywords like “Transparency”) is sufficient, otherwise for dynamic content (like an interactive transparency report dashboard), we can use a headless browser controlled via scripts (eg, Selenium) to render the page and simulate user clicks.[64] This ensures that even content which only appears after the JavaScript code has been run is analysed.[65]

The scrapers respect websites’ robots.txt rules when applicable and throttle requests to avoid overloading servers, legally, when platforms offer official APIs or data feeds for certain information, we utilise those instead, since they are more stable and sanctioned by the provider (eg, some platforms might have an API for their ad library). Robustness is addressed by including into the design of the tool an error handling for common issues (like pages not found, timeouts, or content changes) and logging each step so that a human can review if something goes wrong. The output of this phase is designed to be a structured set of data ready for analysis, for example, the text of the latest terms of service, the list of contact points found, the content of a transparency report PDF, etc.

3.5 Deontic logic formalisation

A distinctive aspect of our approach is the use of deontic logic to model legal rules. Deontic logic is a method for representing normative concepts like obligations (things that must be done) and prohibitions (things that must not be done) in a formal, logical structure.[66]

We created what we call “deontic tables” for each DSA article in our scope. These tables break down an article into individual requirements and link each requirement to measurable criteria and verification methods. For example, for Article 11 and 12’s user contact obligation, one the deontic table would be:

Obligation: Provide an easily accessible contact method for users; Computational Metric: Contact page reachable in ≤2 clicks; Verification: crawl site navigation and count clicks (see Section 4.1 for this case). By doing this for each identified obligation or

[64] ibid.

[65] Daniel Glez-Peña, Analia Lourenço, Hugo López-Fernández, Miguel Reboiro-Jato and Florentino Fdez-Riverola, ‘Web scraping technologies in an API world’ (2014) 15(5) Briefings in Bioinformatics 788.

[66] Roel Wieringa, John-Jules Meyer, Hans Weigand, ‘Specifying dynamic and deontic integrity constraints’ (1989) 4(2) Data & Knowledge Engineering 157, 189.

prohibition, we essentially translate the legal text into a checklist that software can easily follow.[67]

Table 1. Deontic Table Art 11-12

Deontic Norm	Description	Computational Metric	Automated Verification
OB: Easily Accessible Information	The contact point must be clearly identifiable and reachable with minimal steps.	Explicit text with clear contact details, reachable in ≤2 clicks.	Web scraping, text analysis with NLP.
OB: Presence of at Least One Non-Automated Channel	A human contact option must be available.	Identification of an email or phone number among the listed channels.	NER for entity recognition (email, phone number).
IM: Vague or Ambiguous Language	Generic information without precise instructions must not be present.	Detection of vague phrases or generic terms.	NLP models for semantic analysis (BERT, GPT).
OB: Multilingualism	Information must be available in the official languages of the EU.	Number of supported languages compared to regulatory requirements.	Language detection to identify present languages.

The benefit of formalising rules is twofold because it reduces ambiguity in interpretation, and it creates a template that can be consistently applied across many platforms. If the rule says, “provide an email or electronic form for contact”, our formal model might specify “an email address or web form must be present on the contact page.” This removes uncertainty about what counts as compliance, it also allows our system to scan multiple platforms and apply the exact same detection pattern for an email address on each, yielding objective and repeatable comparisons, which can be rigorously explained and verified by humans. In building these tables, we referenced not only the

[67] Dagfinn Føllesdal, Risto Hilpinen, “Deontic Logic: An Introduction” in Risto Hilpinen (ed), *Deontic Logic: Introductory and Systematic Readings* (Springer Dordrecht 1970) 1.

DSA text but also any guidance around it to capture the intent, each row in a deontic table corresponds to one compliance question (eg, “Is there an email address for user contact?”) and the expected answer if compliant (“Yes, at least one email found”). This logical decomposition and its embedding represent the link between law and code in our framework, ensuring that when our system flags something, we can trace it back to a specific legal requirement.

However, it is important to acknowledge that deontic logic, while valuable for formalising normative requirements, cannot fully resolve the inherent ambiguity of legal concepts. Legal interpretation often requires contextual, purposive, and teleological reasoning that exceeds the capabilities of formal logic systems; our design methodology mitigates this limitation by focusing on transparency obligations, which are relatively objective and externally verifiable. Nevertheless, extending this approach to more subjective DSA requirements would necessitate additional interpretative frameworks and more “human-in-the-loop”[68] review mechanisms.

3.6 Automated verification via NLP

Once data is collected and rules are defined, the core analysis occurs. We employ a combination of rule-based checks and Natural Language Processing (NLP) techniques to examine the content for compliance.[69] NLP is crucial because many obligations (like clarity of language or presence of certain statements) involve interpreting text, for instance, determining if terms of service are “easily understandable” is partly subjective, but NLP can assist by measuring reading complexity (eg, average sentence length, use of common vs. legal terms) and by identifying potentially problematic clauses.[70] We designed our system to recognise certain patterns, as for example, detection of contact info (using regular expressions for emails, phone numbers), classification of text as human-oriented or auto-response (using keywords like “no-reply” addresses or the presence of chatbots), of vague language (phrases like “at our sole discretion” might undermine clarity), and identification of multiple languages in a document (through language identification libraries).

For structured data such as transparency report numbers, simple comparisons are done (eg, “does the report include a figure for government removal orders?”). For unstructured text, more advanced NLP, potentially using transformer-based models, helps in semantic analysis. For example, we are developing an NLP classifier that can read a snippet of terms of service and decide if it’s informing users about changes in policy or not. Another NLP

[68] Fabio Massimo Zanzotto, 'Human-in-the-loop artificial intelligence' (2019) 64 Journal of Artificial Intelligence Research 243.

[69] Vijayaragavan Pichiyan, S Muthulingam, G Sathar, Sunanda Nalajala, Ch Akhil, Manmath Nath Das, ‘Web scraping using natural language processing: exploiting unstructured text for data extraction and analysis’ (2023) 230 Procedia Computer Science193.

[70] Lippi and others (n 55).

task is Named Entity Recognition (NER), which we used to differentiate between different types of contact info listed (to ensure human contact is among them, the system must tell an email apart from a chatbot link). The verification module essentially cross-references the scraped data with the computational metrics from the deontic tables; each metric becomes a test (pass/fail or a score) and the module aggregates these to determine overall compliance status for each article per platform. It is important to note that the NLP models chosen in this framework, including rule-based classifiers and transformer-based models, are promising, but these models should be viewed as experimental tools rather than definitive solutions. In fact, their accuracy and robustness in complex legal language processing remain subjects for further empirical validation, particularly in handling nuanced or borderline cases of compliance.[71] Ongoing development focuses on expanding training datasets and integrating domain-specific language models to improve interpretability and reliability. Every automated decision must be systematically logged, and evidence of compliance or non-compliance recorded; this mechanism also must respond to fundamental legal requirements of accountability and transparency, as codified in Article 5(2) GDPR.[72]

Furthermore, it is important to remember that use of artificial intelligence technologies such as those just described, raises important considerations under the AI Act as well, if employed by public authorities for enforcement purposes, even under human supervision and without any degree of autonomy, such tools may nonetheless qualify as high-risk systems insofar as they are intended to assess compliance with legal obligations.[73]

Accordingly, the tool whose design is being proposed must incorporate specific technical and organisational safeguards concerning transparency, documentation, auditability, and human oversight, in line with Articles 9-15 of the Regulation. Compliance with these requirements would not only render the tool legally admissible but would also enhance its reliability and effectiveness as a mechanism for supporting administrative enforcement.

To try to give a more practical description then we highlight how, if the system identifies a breach of Article 12 DSA, it automatically stores the relevant web page content and highlights, for example, that only a chatbot was available. This practice also satisfies the guarantees of effective remedy under Article 47 of the Charter of Fundamental Rights of the European Union (CFREU), enabling supervisory authorities which use the tool to verify and contest enforcement actions based on objective evidence. Moreover, this traceability framework directly refers to auditability obligations established by the DSA, indeed, under Article 15 DSA, providers must keep detailed records of content moderation

[71] ibid.

[72] Elena Gil González and Paul De Hert, 'Understanding the legal provisions that allow processing and profiling of personal data—an analysis of GDPR provisions and principles' (2019) 19(4) ERA Forum 597.

[73] Delaram Golpayegani, Harshvardhan J Pandit and D Lewis, 'To be high-risk, or not to be—semantic specifications and implications of the AI Act's high-risk AI applications and harmonised standards' in *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (2023) 905.

and enforcement decisions, while Article 42 mandates transparency reporting obligations. By facilitating systematic evidence collection and enabling human oversight, the system also respects the safeguards against purely automated decision-making enshrined in Article 22 GDPR, applied here by analogy.

Such a system wants to ensure that regulatory supervision remain dynamic and adaptive, as platforms continually evolve and introduce novel compliance strategies, the systematic recording of evidence allows Digital Services Coordinators[74] and the European Board for Digital Services (Articles 61-63 DSA) to have socio-technical support in analysing patterns of non-compliance, identify systemic risks, and refine enforcement methodologies in line with the cooperation mechanisms outlined in Articles 57 and 58 DSA.

3.7 Data management and interface

All results and data are stored in an organised manner so they can be reviewed and analysed: we use a database management system to keep records of each platform's fetched data (such as a copy of the terms of service text, timestamped) and the outcomes of each compliance check. This historical database allows tracking changes over time for example, if a platform initially failed to provide an ad repository but added one later, we can document that evolution. Historical analysis can reveal trends, such as improvement after regulatory guidance or repeated lapses.

The tool will provide, also, a simple user interface for the enforcement system's output, the interface can be an internal dashboard for regulators where they see, for each platform, a compliance report card: which obligations are met, which are unmet, with details and evidence. There could also be a public-facing component to fulfil the DSA's ethos of transparency, perhaps an anonymised or aggregated view that shows overall industry compliance levels. In our designed prototype, we include features like filtering (to see all platforms that failed a particular requirement) and drill-down (to inspect what exactly was found on a given site). For instance, an enforcement officer could click on "Platform X – Article 11 compliance: FAIL" and see the snippet of the homepage where the contact link was supposed to be, but perhaps wasn't present. The idea is to make the tool's findings accessible and actionable to human decision-makers. By providing a clear presentation, we strive to enable regulators to efficiently focus on the most severe or persistent violations identified by the system.

The methodology combines web scraping for data acquisition[75]), deontic logic modelling for rule formalisation[76], and NLP/data analysis for automated verification. These are supported by a data management backend and an interface for results. Through these

[74] cf Article 49 DSA.

[75] Khder (n 62).

[76] Wieringa and others (n 65).

components, we translate the DSA's requirements into an automated workflow that can monitor numerous platforms continuously.

At the current stage, the proposed design is being developed as a prototype, primarily designed for experimental validation and proof-of-concept demonstrations. While the system effectively automates core compliance checks across selected DSA provisions, it remains an experimental tool requiring further refinement before large-scale deployment, future development plans include testing on a wider set of platforms and integration with official regulatory data sources.

The next section will show how this framework is applied to the specific DSA articles we selected, detailing what the system can verify for each article and what findings it can produce.

4 Case studies: automated enforcement for selected DSA obligations

4.1 Contact point obligations (Articles 11-12)

Articles 11 and 12 of the DSA ensure that communication channels exist between platforms, regulators, and users. Article 11 requires intermediaries to have a single point of contact for authorities (eg, a dedicated email for officials to send takedown orders or inquiries). Article 12 requires an easily accessible electronic contact method for users to reach the platform, and crucially, this contact method cannot rely solely on automation (so users shouldn't be forced to talk to a bot with no option of human support). Using our socio-technical framework, we are automating the verification of these requirements as follows.

For each platform in our sample, the system navigates to find any "Contact Us" or legal notice page. It checks if an official contact point for authorities is listed (some platforms have a section like "Law Enforcement Inquiries"). At minimum, Article 11 compliance might be evidenced by a statement such as "Regulatory authorities may contact us at legal@platform.com." Our scraper searches for keywords like "authority" or "DSA" on relevant pages. If none are found, it flags that Article 11 may be unmet. For user contact (Article 12), the system looks for a general contact email or form accessible to ordinary users, it evaluates accessibility by seeing how many clicks it takes from the homepage to reach that information. We defined a threshold (two clicks) as a reasonable measure of "easily accessible," based on web usability norms and the idea that users shouldn't have to dig through many pages.[77] If our crawler had to traverse an overly complicated path, that's recorded as a potential violation.

To address the "not exclusively automated" clause, our NLP component analyses the text around the contact method. If the only contact offered is a chatbot or a list of FAQs

[77] Section 3.2 discussed this metric.

(frequently asked questions) without any direct human email or phone, we mark this as non-compliant. We trained a simple classifier to differentiate between contact info that likely reaches a human (eg, presence of an email address or a physical office address) and purely automated channels (eg, a link that says "Chat with our virtual assistant"). The presence of an email address or a support ticket form that promises human follow-up satisfies the requirement.

Additionally, we check if the contact page content is provided in multiple languages (since DSA encourages platforms operating in the EU to cater to different official languages for user-facing information). Through language detection on the contact page, the tool can note if, for example, a platform only provides contact info in English despite operating in several EU Member States.

4.2 Terms of service (Article 14)

Article 14 obliges platforms to be transparent and fair in their terms of conditions (ToC), especially regarding content moderation rules. It also requires notifying users of significant changes and includes special considerations if a service is widely used by minors. Our automated approach treats the platform's Terms of Service document as the primary data source to assess compliance with respect to Article 14.

Firstly, the scraper retrieves the latest Terms of Conditions or User Agreement page, indeed, our system is designed to scan it for certain content moderation policy disclosures that Article 14 expects. For instance, the DSA requires that terms clearly explain any rules about permissible content or user behaviour (so users know what could lead to removal or suspension). Using keyword searches, we identify if the terms mention things like "we may remove content that … (violates X)" or a section on "Content guidelines" exists, if such sections are missing or very unclear, that's a red flag.

Next, we evaluate the clarity and accessibility of the terms, and the system calculates the Flesch-Kincaid reading ease score.[78] While legal terms are often complex, an extremely difficult score might indicate the text is not "clear and unambiguous" as the DSA intends, after computing scores for a statistically significant sample of ToCs, a warning and a critical threshold will be provided in our deontic tables to guarantee grounded and reproducible results.

We also look for formatting that aids understanding, such as headings, summaries or bullet points. Some forward-thinking platforms include a summary or FAQ alongside their full legal terms. If our tool finds a summary (for example, a TL;DR section), it notes that as a positive compliance feature aligned with accessibility best practices. Conversely, if the entire ToS is a single dense block of text, we highlight that as problematic from an accessibility standpoint.

[78] J Peter Kincaid, Robert P Fishburne Jr, Richard L Rogers, Brad S Chissom, 'Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel' (Millington 1975).

To check compliance with notification of changes, our design leverages a version history, indeed, we store the fetched terms of service text with a timestamp. When run periodically, the system can compare the new version to the old, if changes are detected and Article 14 requires that users be informed of "any significant change," we look on the platform's site for evidence of such notice (often platforms will post a blog update or a banner announcing updated terms). If our periodic check finds that terms have changed but no announcement was found, that could indicate non-compliance. In testing, we found one instance where a platform updated its terms (the text differed) but the only way a user could know was by checking a small date stamp on the terms page-arguably not a sufficient notice, this would be flagged for regulators to examine.

We also incorporate the minor protection aspect: if a platform is known to be popular with minors (for example, a gaming or social media app), Article 14 expects the terms to be appropriate for that audience. Our tool doesn't have an age-popularity database built in, but we used an external list of youth-oriented services to trigger this check. For those services, we ensure the terms include any special provisions for under-18 users.[79]

Overall, the automated analysis of ToS produces a multi-faceted result: a measure of readability, a checklist of required disclosures (found or not), an indicator of whether changes are being tracked and communicated, and notes on any unusual or potentially unfair clauses (our NLP flags extremely one-sided clauses like "we can remove content for any or no reason" as something that might undermine transparency). While the system cannot judge fairness in a legal sense, it points out elements that human regulators or courts might scrutinise under the DSA's provisions against unfair terms. For this reason, another future work consists in defining a total compliance score for Terms of Service obtained by computing sub-scores for each compliance check defined in this subsection, following the example set by the Flesch-Kincaid reading ease score, and aggregating them by calculating the weighted sum using appropriate coefficients, which will be established after an in-depth evaluation of a statistically significant sample of ToS documents.

4.3 Transparency reporting (Article 15)

For Article 15, the input data is the platform's transparency reports. Many large platforms publish periodic reports (quarterly or biannually) detailing metrics such as how many pieces of content were removed, how many user notices were received, average response times, etc, as required by the DSA. Our automated tool is configured to fetch these reports (often PDFs or web pages).

We parse the content of each report and verify that it contains certain core statistics mandated by the DSA: for example, number of orders from public authorities to remove content, number of content removal actions taken by the platform on its own initiative,

[79] for example, parental consent clauses or simplified language sections.

breakdown of the reasons for removals, number of complaints received and processed, and outcomes of those complaints. We have templates of expected sections, and the system searches the text for corresponding keywords (like "government requests: X" or "content removed: Y"). If any key metric is missing, that's a compliance issue.[80]

Additionally, we designed a way to examine the granularity and format. Article 15 expects that the information be provided in a way that allows understanding and analysis. If a report is extremely high-level or aggregates things too broadly, it might not fulfil the obligation. For example, if a platform simply states "We removed 50,000 posts last year" without context or category, it's not very transparent. Our method doesn't fully judge the sufficiency of detail (which can be subjective), but it does compare the report content against known standards or typical reports from peers. A platform that provides a multi-page detailed report will pass our checks easily, whereas one that posts a one-paragraph summary will likely fail some checks (like missing breakdowns per category of illegal content).

We also aim to integrate a timing check; indeed, Article 15 requires reports at least once a year. Our system notes the dates of the reports and can alert if a scheduled report is overdue or if the interval is too long. For instance, if a VLOP hasn't updated its transparency report in over a year, it's likely not compliant and the system will show it.

4.4 Notice-and-action mechanism (Article 16)

Article 16 ensures users have a channel to notify platforms of illegal content and receive a timely response, which is central to user empowerment in content moderation, for this, our method for evaluating Article 16 is somewhat interactive because we simulate the role of a user trying to report content. While we do not actually submit reports (to avoid sending false reports to platforms), we go through the motions up to the point of submission.

The system checks if each piece of user-generated content (eg, a post or video) on the platform has an obvious "Report" function, we are training it on known patterns (a flag icon, a "Report" button in dropdown menus, etc.) and if none is found on a representative sample of content, that's a direct violation (users have no way to report).

If a reporting interface exists, our tool accesses it (for example, clicking "Report" opens a form or modal); we then analyse the options provided within the form, including whether it allows the user to specify the type of illegal content (such as hate speech, piracy, etc) in accordance with the expectations set by the DSA. We also examine whether the process is accompanied by explanatory information, such as statements like "We will review and respond within 24 hours" or the presence of a confirmation message.

[80] Reid, Ringel (n 47).

Our system also is designed to look for terms of service or help centre descriptions of the notice-and-action procedure, indeed, Article 16 requires platforms to acknowledge receipt of notices and inform users of decisions. We search the site for statements like "you will receive an email confirmation" or "we will inform you of action taken." If the platform publicly describes such a process, we take that as a sign of compliance (and if the description is absent, it might indicate the process isn't well established).

We also consider usability factors like if there is a reporting mechanism hidden behind too many clicks or one that requires unwarranted information from the user could be non-compliant because it's not "easily accessible." Our designed tool aims to time how long it takes to reach the final stage of the report form, if it's overly convoluted or, say, only available to logged-in users when it should be open, it notes that.

Since full testing would involve submitting actual notices and awaiting platform responses (which is beyond our automated script's ethical scope), we flag elements that suggest whether the follow-through happens, for instance, if the platform's form asks for the user's email, that implies they will send a confirmation (positive sign), if it doesn't that discrepancy is flagged.

Through these steps, the tool can output an assessment like: "Platform X provides a reporting form reachable through two clicks on each post. The form covers the required categories of illegal content and promises a confirmation email (as evidenced in help pages). Platform Y, however, only allows reporting via a generic contact email found in the help section, which is less accessible and provides no info on response times potentially not fulfilling Article 16 requirements." This kind of comparative, automated review helps regulators quickly see which platforms might be making it hard for users to report issues, thereby undermining the DSA's notice-and-action system.

4.5 Advertising and recommender transparency (Articles 26-27)

For VLOPs and other large services, advertising and recommendation disclosures are novel obligations that our methodology tackles by a mix of content scraping and interface inspection.

To check ad transparency (Article 26), our system does two main things: it looks at the interface where ads appear to ensure they are labelled, and it searches for the platform's ads repository. For labelling, the scraper might load a user feed and identify sponsored content elements (most platforms embed a label like "Sponsored" or "Ad" in the HTML), if our parser finds posts that seem to be ads (by structure) but without a clear label in the text, that's a failure to meet the basic transparency of labelling ads; then, for each detected ad, we try to find the "Why am I seeing this ad?" feature (common on Facebook, Twitter, etc).[81] That usually brings up a pop-up or page with details on targeting, we

[81] Tami Kim, Kate Barasz and Leslie K John, 'Why am I seeing this ad? The effect of ad transparency on ad effectiveness' (2019) 45(5) Journal of Consumer Research 906.

capture that content and check if it includes the required information: advertiser identity and targeting criteria; if a platform does not provide that detail or such a feature is missing, it likely violates Article 26's second part. Additionally, the DSA's requirement of an ad repository means there should be a publicly accessible archive of ads, our tool attempts to find this by looking for links titled "Ad Library" or scanning the sitemap/robots file for references to an ads archive, if found, it can scrape it to see if it's functional (though analysing its completeness is complex, we at least verify it exists and is reachable), indeed, a missing ad repository for a platform that should have one is a significant compliance gap.

For recommender systems transparency (Article 27), the checks are somewhat qualitative; we search for a user-accessible explanation of the main parameters. Platforms often implement this via a "Personalisation settings" page or an info box that explains, for example, "Your feed is sorted by relevance, which takes into account your likes and follows." We gather such text and evaluate clarity (is it in plain language?) and completeness (does it mention key factors like user behaviour, popularity, etc?). We also verify the presence of a toggle or option for users to adjust recommender settings. The DSA effectively gives users the right to influence how content is recommended, which many interpret as offering at least a chronologically sorted feed or some non-personalised alternative. Our proposed automated test looks in the settings menu for any option related to feed order or recommendations: if not found, we suspect non-compliance.

One challenge with Article 27 is that simply stating "our algorithm suggests content based on your interests" might be technically compliant but not very useful. We leverage some criteria from emerging best practices on algorithmic transparency to gauge depth. For example, we consider it a better compliance if the platform enumerates specific input signals (like "we use your watch history and your location") rather than vague statements. Finally, the system notes if user controls are effective. We can test a simple scenario: if a user opts out of personalised recommendations (if that option exists), does the feed change order or content? This is tricky to do automatically, but we can at least confirm if such an option triggers any visible change in the HTML or if the platform acknowledges the choice ("You are now seeing posts in chronological order"). A completely static response might indicate the option is decorative rather than functional.

In applying our automation to Articles 26–27 on a sample platform, we might get results like: "Platform X clearly labels ads and provides an accessible ad library link (compliant with Art 26). It also gives users the choice between a personalised and chronological feed and explains in its help centre that recommendations are based on user activity (mostly compliant with Art 27, though explanation could be more detailed). Platform Y, however, does not visibly label ads: our scraper could not find any 'Sponsored' tags on ads, and we could not locate any public ad archive. Its feed is algorithmic with no user toggle, and no explanation of how content is chosen was found, this suggests Platform Y falls short on both ad and recommender transparency requirements." Such findings underscore the areas where automated tools can immediately highlight likely non-compliance, prompting enforcement action or further inquiry.

5 Conclusion

This study has presented a socio-technical framework that combines legal analysis with technical innovation to help automate the enforcement of the DSA's transparency obligations. In doing so, it provides a view of how technology can complement traditional regulatory oversight.[82]

Our proposed framework, centred on web scraping, natural language processing, and logical formalisation of rules, wants to offer to Authorities a scalable tool to monitor whether digital platforms are meeting their DSA duties in real time. By systematically checking for contact points, scanning terms of service for clarity, validating transparency reports, simulating user notice processes, and inspecting interfaces for ad and algorithm disclosures, the approach translates high-level legal requirements into actionable audit tasks that a computer can perform across many services at once.

The advantages of such an automated enforcement tool are evident in a landscape where manual supervision is increasingly impractical. Platforms generate enormous amounts of data, and their practices evolve rapidly, a human-only enforcement regime would struggle to keep up.[83]

Automation improves speed and consistency so it can quickly identify issues when a platform deviates from compliance, and it applies the same standards uniformly, reducing the risk of oversight being uneven or biased, therefore enhancing the effectiveness and credibility of enforcement. Moreover, by providing a continuous check, it encourages platforms to maintain compliance proactively (knowing that lapses will be caught sooner than later), thereby furthering the DSA's goals of accountability and user protection.

However, it is important to acknowledge the limitations and challenges that remain, firstly, the system's assessments are only as good as the rules and patterns it's given; complex legal interpretations or context-specific judgments are still difficult to encode. There's a risk of false positives (flagging non-issues) or false negatives (missing subtle forms of non-compliance) if the logic isn't carefully calibrated. For example, an overly strict parser might mark legal language as "unclear" when it's acceptable or miss a cleverly hidden contact link. We are addressing this by proposing the incorporation of human reviewers in the loop as our results are meant to aid, not replace, human regulators.[84] The framework provides leads and evidence, but enforcement decisions will often require a human confirming that a violation is real and significant. This hybrid model

[82] Hassan, De Filippi (n 12).

[83] Afzal (n 30).

[84] Sriraam Natarajan, Saurabh Mathur, Sahil Sidheekh, Wolfgang Stammer and Kristian Kersting, 'Human-in-the-loop or AI-in-the-loop? Automate or collaborate?' in *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol 39, no 27 (2025) 28594.

is likely to persist; full automation without oversight could lead to disputes, especially if a platform argues the tool misinterpreted something (which is why our design logs evidence for transparency).

Secondly, technical improvements are needed to keep the enforcement tool robust, indeed, platforms may change their site structure or even attempt to game automated checkers (in a scenario where they know regulators use them). Ongoing development of more sophisticated NLP models that understand context and nuance (eg, distinguishing a genuine attempt at clarity from legalese) will help: advances in AI, such as domain-specific language models (like a "LegalBERT" trained on policy documents), could enhance the tool's ability to interpret terms of service and other text with greater fidelity to legal meaning. We foresee integrating such models to better evaluate qualitative aspects, like fairness of terms or adequacy of algorithmic explanations, which are areas current simple checks only approximate. Future works will explore in-depth each component employed in the framework analysed in this paper by implementing and evaluating experiments aimed at providing numerical and replicable results.

Another consideration is ensuring the enforcement tool itself is transparent and accountable, just as we demand platforms to be transparent, any regulatory algorithm should be explainable. We have built in explainability by using deontic logic tables that clearly map to legal provisions and by recording how conclusions are reached. As this project advances and is implemented, we would aim to further open the tool's methodology, perhaps even publishing an open dashboard showing overall compliance statistics across platforms (without revealing confidential details). This could empower civil society and researchers to participate in oversight, aligning with the DSA's aim to foster a public-private enforcement.

Our paper concludes that automated enforcement can enhance the effectiveness of the DSA, assuming that automated tools are carefully designed and implemented by private business and public authorities. Regulators would need training to use such tools and processes to respond swiftly to the findings (eg, sending notices to platforms when an issue is flagged or coordinating between EU countries if widespread non-compliance is detected). The tool could also benefit from input by platforms themselves; for instance, if platforms share data or APIs for compliance info, the tool can plug in to get more reliable data than scraping. Encouraging a cooperative approach where platforms are aware of the automated checks and perhaps even pre-emptively use similar tools internally to audit their compliance could create a constructive compliance culture.

More generally, the approach we have developed can serve as a framework for regulatory automation in other domains, the concept of translating legal obligations into machine-readable rules and verifying them through data analysis could apply to data protection (imagine a tool scanning a website for GDPR compliance indicators), consumer protection (automatically detecting unfair clauses in terms[85]), or financial services compliance, among others. As digital regulation expands, regulators will increasingly need tech-assisted methods to uphold the law effectively. Our work contributes to that emerging field of RegTech (regulatory technology) by showing a practical example in the realm of platform governance.

In conclusion, the enforcement of the DSA stands to benefit greatly from the integration of automated, intelligent systems. Such systems can ensure that transparency, accountability, and user rights, the very values the DSA champions, are not lost in the immense scale of the online ecosystem. By operationalising legal requirements through code, we take a step toward a future where regulation is written in law but also embedded in the digital infrastructure of platforms. This cross-disciplinary effort, uniting law and technology, aims to uphold democratic values in online spaces in a consistent and timely manner. While challenges remain and continual refinement is needed, the design outlined in this article offers a perspective for enhancing DSA enforcement and, ultimately, for fostering a safer and more transparent digital world.

[85] Lippi and others (n 55).