European Journal of Risk Regulation (2026), 1–14
doi:10.1017/err.2026.10119

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Turning the DSA into a Verification Tool: A Feasible Solution to the Reliability Problem

Alessia Zornetta

Institute for Technology, Law & Policy, University of California Los Angeles, USA
Email: alessiazornetta@gmail.com

Abstract

Content-moderation transparency suffers from a fundamental reliability problem: platforms disclose information about their practices, but the accuracy of those disclosures is difficult to verify. The Digital Services Act (DSA) introduced a Database of Statements of Reasons that has collected billions of decision-level records since February 2024, and, since July 2025, the Implementing Regulation has harmonised transparency-reporting templates and aligned categorisation across instruments. This article assesses whether those reforms convert the Database from an observation tool into a verification tool. The article argues that while the reforms deliver semantic alignment and comparability, they leave the legal relationship between aggregate reports (Articles 15, 24) and decision-level data (Article 17) undefined. A narrowly tailored duty of traceability – linking reported aggregates to the corresponding statements of reasons – would enable systematic consistency checks without expanding the scope of reporting. Absent this clarification, transparency will remain descriptive rather than verifiable, particularly for intermediaries outside the audit regime.

Keywords: Digital Services Act; European Union Law; platform law; technology law; transparency

I. Introduction

Transparency has long been regarded as a cornerstone of accountability and trust in both democratic governance and corporate regulation. Rooted in the *right to know*, it empowers citizens and consumers to scrutinise decision-making and ensure compliance when self-regulation fails.[1] In the digital age, transparency has become indispensable to platform accountability, as content moderation profoundly shapes online discourse yet remains opaque.[2] Public, academic, and regulatory calls for transparency have remained constant in platform governance but their lack of specificity has also enabled platforms to frame selective disclosures as compliance tools rather than meaningful oversight mechanisms.[3] To become a

[1] A. Fung, "Infotopia: Unleashing the Democratic Power of Transparency" (2013) 41 Politics & Society <https://doi.org/10.1177/0032329213483107> (last accessed 8 January 2023).

[2] N. P. Suzor and others, "What Do We Mean When We Talk About Transparency? Toward Meaningful Transparency in Commercial Content Moderation" (2019) 13 International Journal of Communication 18, 1527.

[3] A. Zornetta, "Clearing the Haze: Is the EU Digital Services Act Finally Forcing Platforms to Open Up About Content?" (2025) 6 Notre Dame JET 5; Suzor and others (n 2).

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial licence (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original article is properly cited. The written permission of Cambridge University Press or the rights holder(s) must be obtained prior to any commercial use.

https://doi.org/10.1017/err.2026.10119

vector for accountability, transparency must be meaningful: receivers of transparency must be able to interpret, understand, and act upon the information disclosed.[4] Yet, even when transparency mechanisms are in place, their effectiveness depends on the reliability of the information they disclose. In the context of platform governance, the central challenge is not transparency *per se*, but verifiable, meaningful transparency that allows independent confirmation of disclosed practices rather than mere observation.

Existing disclosure practices largely enable platforms to describe their content moderation systems while leaving external actors without tools to assess their accuracy. The transparency requirements in the Digital Services Act (DSA) were designed to address precisely this issue. Conceived as a "transparency machine," the DSA sought to unveil the opaque practices of content moderation that had long avoided public scrutiny and to position itself as a cornerstone in the emerging architecture of platform governance.[5]

Yet, the first transparency reports fell short of these expectations. The absence of a standardised reporting framework produced disclosures that were inconsistent, difficult to interpret, and ultimately incapable of realising the DSA's ambition of generating meaningful transparency.[6] More importantly, the framework lacked mechanisms to verify the accuracy and completeness of the information provided.[7] Indeed, the DSA's auditing obligations are limited to Very Large Online Platforms and Search Engines (VLOPs/VLOSEs) and focus primarily on procedural adequacy rather than the substantive accuracy of reported data.[8] For all other intermediaries, verification rests with oversight authorities empowered to request additional information or initiate investigations but the practical feasibility of exercising such powers across hundreds of intermediary services is doubtful.[9]

Resource constraints, jurisdictional fragmentation, and competing enforcement priorities make comprehensive verification of platforms' disclosures virtually unattainable. This structural limitation effectively preserves platforms' discretion to curate their self-reporting, thereby undermining the reliability of the transparency regime.[10] Paradoxically, the DSA already provides the technical infrastructure necessary for verification. The Database of Statements of Reasons contains decision-level data that, if linked to the aggregate figures disclosed by platforms under Articles 15 and 24, could enable systematic consistency checks and independent validation.[11] The 2025 harmonisation reforms[12] to the transparency reporting templates and the Database schema could

[4] Ibid. *supra* note 3, 5–6; D. Keller and P. Leerssen, "Facts and Where to Find Them: Empirical Research on Internet Platforms and Content Moderation" *Social Media and Democracy: The State of the Field and Prospects for Reform*, vol 220 (Cambridge, UK, Cambridge University Press 2020) p 224.

[5] Zornetta, *supra* note 3, 37.

[6] Zornetta, *supra* note 3.

[7] B. Wagner and others, "Auditing Big Tech: Combating Disinformation with Reliable Transparency" (2021) Enabling Digital Rights and Governance 4. (arguing that "without being able to verify platform's data, existing public debate is based merely on a whim, and digital dominance is likely to become more pronounced.").

[8] DSA Article 37. Even in this case, the effectiveness of DSA audits may be limited. See J. Albert, "DSA Risk Assessment Reports: A Guide to the First Rollout and What's next – DSA Observatory" (9 December 2024) <https://dsa-observatory.eu/2024/12/09/dsa-risk-assessment-reports-are-in-a-guide-to-the-first-rollout-and-whats-next/> (last accessed 10 February 2025); A A. Zornetta, "Beyond Self-Reporting: Verifying Platform Transparency Under the EU Digital Services Act" (2026) Rutgers Law Record.

[9] S. Demková and G. De Gregorio, "The Looming Enforcement Crisis in European Digital Policy" (2025) Verfassungsblog <https://doi.org/10.59704/c4fc2ec9cb2b9ba5> (last accessed 6 November 2025).

[10] See G. K. Shahi and others, "A Year of the DSA Transparency Database: What It (Does Not) Reveal About Platform Moderation During the 2024 European Parliament Election" (arXiv, 9 April 2025) <https://doi.org/10.48550/arXiv.2504.06976> (last accessed 6 November 2025).

[11] A. Zornetta, "The DSA Database as a Tool for Transparency Accountability" (2025) 2025 University of Illinois Journal of Law Technology and Policy <https://doi.org/tra>.

[12] Commission Implementing Regulation (EU) 2024/2835 of 4 November 2024 laying down templates concerning the transparency reporting obligations of providers of intermediary services and of providers of

https://doi.org/10.1017/err.2026.10119

have facilitated such alignment, shifting transparency from a self-reporting exercise to a verifiable system of accountability. Whether these reforms will succeed in converting transparency from descriptive to evidentiary is the central focus of this article.

Part II will discuss how the initial framework for transparency requirements under the DSA perpetuated the reliability problem in content moderation transparency reporting, whereby platforms provide an illusion of transparency by regularly publishing information about their practices, but avoid accountability by preventing independent verification of their disclosures. Part III analyses the Implementing Regulation and the harmonised Database of Statements of Reasons schema to assess whether the reforms have addressed the reliability problem identified in Part II. Part IV discusses the under-utilisation of the regulatory tools in the DSA transparency schema and argues that it stems not from its technical incapacity but from the absence of regulatory design for verification, ultimately identifying the legal, technical and institutional preconditions required for the Database to evolve into a verification tool capable of reinforcing the reliability of the DSA transparency framework as a whole. Part V concludes.

II. The initial DSA framework and its reliability problem

Before assessing the reforms introduced by the 2025 Implementing Regulation[13], it is necessary to examine the original structure of the DSA's transparency regime and the reasons why it failed to produce verifiable information. The DSA's first implementation cycle (2023–2024) dramatically expanded the amount of data that intermediary services were required to disclose but left unresolved the central question of reliability. This section argues that the DSA's architecture combined extensive reporting duties with narrowly circumscribed verification mechanisms, resulting in a transparency system that generated visibility without accountability.

1. Mandatory transparency obligations under the DSA

The DSA introduced a graduated framework of transparency obligations that scales in intensity with a provider's size, function, and societal impact.[14] Articles 15, 24 and 42 set out the core reporting requirements, while Articles 37 and 40 create limited avenues for verification through audits and researcher access.

At a minimum, all providers of intermediary services must publish annual transparency reports detailing orders from Member States' authorities, notices submitted under the Article 16 notice-and-action mechanism, own-initiative content moderation actions, complaints received through internal complaint-handling systems, and the use of automated tools. For all these categories, platforms must include details on type of allegedly infringing content, action taken, median time needed to react, as well as use of automated systems.

Online platforms face additional obligations under Article 24, including the disclosure of user-base metrics and outcomes of out-of-court dispute settlements.[15] VLOPs and VLOSEs must report every six months and include further detail on their moderation workforce, including linguistic expertise, training protocols, and indicators of accuracy.[16]

online platforms under Regulation (EU) 2022/2065 of the European Parliament and of the Council [2024] OJ L 2024/2835, 5.11.2024, ELI: http://data.europa.eu/eli/reg_impl/2024/2835/oj ("Implementing Regulation").

[13] Ibid.

[14] DSA Recitals 40–1.

[15] DSA, Art. 24.

[16] DSA, Art. 42.

https://doi.org/10.1017/err.2026.10119

Although this architecture increased the quantity and periodicity of data available, it offered few means to verify accuracy and completeness. Only VLOPs and VLOSEs are subject to annual independent audits under Article 37, where auditors must assess procedural compliance. For all other intermediaries, oversight remains confined to the general investigatory powers of the Commission and national Digital Services Coordinators. In practice, these authorities lack the capacity to review hundreds of reports across the EU. The result is an asymmetric system in which the majority of intermediaries publish unverified data while the most powerful actors undergo audits focused primarily on processes rather than outcomes.[17]

While Articles 15 to 24 determine what platforms must report, Article 17 establishes how they must document individual moderation decisions. The interaction – or lack thereof – between these two instruments will prove decisive for the reliability of the overall framework.

2. *The DSA database and transparency-reporting limitations*

Among the DSA's transparency instruments, the Database of Statements of Reasons (hereafter, Database) represents its most innovative and potentially transformative component. Since February 2024, all online intermediaries have been required to log in this centralised repository every "statement of reasons" (SoR) issued to users whose content or accounts has been affected by a content moderation decision. The Database provides, for the first time, insights into individual decisions by recording, for each action, the relevant facts, the legal or contractual basis, the use of automated means during the detection and decision-making processes and the restriction applied.

The Database is governed by articles 17, 24 and 43. Article 17 requires that platforms provide users with a clear and specific statement about the content moderation decision and possibilities for redress. Article 24(5) obliges providers to submit each statement "without undue delay" to the Commission's Database in machine-readable form, and Article 43 establishes the funding system for the Database maintenance. The Database's scope, however, is not exhaustive and excludes deceptive high-volume commercial content, government-ordered removals and law-enforcement-sensitive actions.

Technically, the Database operates through an application programming interface (API) that allows single or bulk submissions and enforces predefined attributes for content type, restriction category, detection method, territorial scope and legal basis, standardising documentation and enabling automated data feeding at scale.[18] Nevertheless, an analysis of the initial design of the Database and the SoR template revealed limitations that constrained its potential as accountability tool.[19] The original Database's taxonomy diverged from platforms' classifications in transparency reports, precluding cross-dataset comparison; redress information required under Article 17(3)(f) was not captured; and entries could not be amended to reflect reversals or appeals. Above all, the 2024 schema included no legal or technical mechanism to link decision-level data with the aggregate figures reported under Articles 15 and 24.

As a result, the initial Database functioned primarily as a descriptive repository rather than a verification instrument. While it greatly expanded visibility into moderation

[17] M. Jóźwiak, "The Wait Is (Almost) over! First Risk Assessment and Audit Reports – What Will Be Published, When, and the Way Forward - DSA Observatory" (22 November 2024) <https://dsa-observatory.eu/2024/11/22/the-wait-is-almost-over-first-risk-assessment-and-audit-reports-what-will-be-published-when-and-the-way-forward/> (last accessed 23 March 2025); Zornetta, *supra* note 8).

[18] "Overview Documentation – DSA Transparency Database" <https://transparency.dsa.ec.europa.eu/page/documentation#the-legal-or-contractual-grounds-relied-on-in-taking-the-decision> (last accessed 29 October 2025).

[19] Zornetta, *supra* note 11.

https://doi.org/10.1017/err.2026.10119

activity, it did not allow regulators or researchers to test the accuracy of platforms' aggregate disclosures. Only with the 2025 harmonisation process did the Commission begin addressing these structural deficiencies – a reform analysed in Section III. The following section delves into how this early disconnect between datasets exemplified and perpetuated the reliability problem inherent in the DSA's first transparency reporting cycle.

3. The reliability problem

The reliability problem arises when transparency practices allow platforms to shape the narrative concerning their content moderation processes while preventing external validation of the claims they advance.[20] The selective structure of transparency reports, where companies determine what to disclose and how to frame it, has long hindered the efficacy of voluntary transparency practices, transforming disclosures into instruments of reputational management that project the appearance of openness while remaining substantively evasive or incomplete.[21] Where legal obligations are absent or weakly enforced, platforms face no material consequences for omissions or inaccuracies, visibility is rewarded and disclosure is treated as an end in itself.[22]

Even after the DSA introduced mandatory reporting, the reliability problem persisted. The DSA sanctions mainly the failure to report, not the quality of reported information, thereby allowing formal compliance to coexist with substantive opacity.[23] In the absence of any systematic verification processes to assess the veracity and completeness of platforms' transparency disclosures, platforms retain discretion to curate what they reveal, omitting or downplaying information that could invite criticism.[24]

Auditing obligations and researcher access to internal data – the few verification elements in the DSA transparency framework – are limited to VLOPs and VLOSEs, and, even there, focused primarily on procedural adequacy rather than factual correctness. For the majority of online platforms, oversight is dependent on authorities constrained by limited resources and competing enforcement priorities. Inaccurate or strategically curated reports may therefore go undetected unless they attract targeted regulatory scrutiny. The result is a transparency regime that expands the quantity of publicly available information while leaving its reliability uncertain and an uncaptured potential. As the first analyses of the DSA Database demonstrate, selective or inconsistent disclosures remain common even under a binding legal framework.[25]

[20] K. Sorensen, "What's Wrong with Transparency Reporting (and How to Fix It)" (ADL 2021) <http.s://www.adl.org/resources/report/whats-wrong-transparency-reporting-and-how-fix-it> (last accessed 8 January 2023).

[21] Ibid; R. Gorwa and T. G. Ash, "Democratic Transparency in the Platform Society" [2020] Social Media and Democracy: The State of the Field, Prospects for Reform 286; A. Urman and M. Makhortykh, "How Transparent Are Transparency Reports? Comparative Analysis of Transparency Reporting across Online Platforms" (2023) 47 Telecommunications Policy 102477 <https://doi.org/10.1016/j.telpol.2022.102477>; A. Reid, E. Ringel and S. M. Pendleton, "Transparency Reports as CSR Reports: Motives, Stakeholders, and Strategies" (2023) 20 Social Responsibility Journal 81 <https://doi.org/10.1108/SRJ-03-2023-0134>.

[22] H. Tworek, "How Transparency Reporting Could Incentivize Irresponsible Content Moderation" (*Centre for International Governance Innovation*) <https://www.cigionline.org/articles/how-transparency-reporting-could-incentivize-irresponsible-content-moderation/> (last accessed 6 November 2025).

[23] Shahi and others, *supra* note 10.

[24] Zornetta, *supra* note 11, 213; L Vigneau and C Adams, "The Failure of Transparency as Self-Regulation" (2023) 14 Sustainability Accounting, Management and Policy Journal <https://doi.org/10.1108/sampj-01-2022-0051> (last accessed 6 November 2025).

[25] Shahi and others, *supra* note 10; A. Trujillo, T. Fagni and S. Cresci, "The DSA Transparency Database: Auditing Self-Reported Moderation Actions by Social Media" <https://doi.org/10.48550/ARXIV.2312.10269> (last accessed 18 September 2024); C. Papaevangelou and F. Votta, "What Platform Observability Have You given Us? A First Look into the Statement of Reasons Database" (DSA and Platform Regulation Conference, Amsterdam

https://doi.org/10.1017/err.2026.10119

Unless transparency mechanisms are coupled with verifiability, disclosure risks reinforcing platforms' narrative control rather than constraining it. It is within this context that the Database of Statements of Reasons assumes particular importance: properly leveraged, it could supply the empirical foundation for cross-checking aggregate reports and thereby transform transparency from a descriptive obligation into an evidentiary one. The next section examines whether the 2025 reforms – through the Implementing Regulation and the harmonised Database schema – have addressed the reliability problem.

III. The Implementing Regulation: reforms and expectations

Adopted on 4 November 2024, the Implementing Regulation laying down templates for transparency reporting (Implementing Regulation) represents the EU's first systematic attempt to standardise how intermediary services disclose information under Articles 15, 24 and 42 DSA.[26] The Commission presented the reform as a "step towards harmonised and comparable reports," seeking to replace the fragmented and heterogeneous disclosures that characterised the DSA's initial implementation cycle.[27]

By introducing common reporting deadlines, an exhaustive taxonomy of categories and uniform data-formatting rules, the Implementing Regulation aims to limit platforms' discretion and render disclosures comparable both across services and over time.[28] These changes also resulted in parallel updates to the template for SoR submissions to the Database, theoretically aligning the granular data stored in the Database with the aggregated figures disclosed in transparency reports. Whether this harmonisation can also render them verifiable, however, remains an open question.

1. Key changes to transparency reporting requirements

The Implementing Regulation introduces a comprehensive set of harmonisation measures designed to transform the DSA's transparency reporting framework from a loose collection of self-defined disclosures into a standardised, machine-readable system. The first reform standardises the temporal scope of transparency reports.[29] This alignment resolves one of the most practical barriers to cross-platform comparability. During the first three reporting cycles, platforms selected divergent reference periods, producing temporal gaps that rendered cross-platform comparison particularly challenging.[30] Reports describing overlapping but non-identical timeframes captured moderation activity from inconsistent intervals, frustrating attempts to establish baselines or identify sector-wide trends.[31] The harmonised reporting periods now ensure that disclosures refer to the same temporal

Law School, 15 February 2024) <https://dsa-observatory.eu/the-dsa-and-platform-regulation-conference-2024/>; C. P. Drolsbach and N. Pröllochs, "Content Moderation on Social Media in the EU: Insights From the DSA Transparency Database" *Companion Proceedings of the ACM Web Conference 2024* (Association for Computing Machinery 2024) <https://doi.org/10.1145/3589335.3651482> (last accessed 9 September 2024).

[26] "Implementing Regulation," *supra* note 12.

[27] "Harmonised Transparency Reporting Rules under the Digital Services Act Now in Effect | Shaping Europe's Digital Future" <https://digital-strategy.ec.europa.eu/en/news/harmonised-transparency-reporting-rules-under-digital-services-act-now-effect> (last accessed 25 October 2025).

[28] Ibid.

[29] Ibid. As of July 2025, intermediary services must submit their annual transparency reports by the end of February each year, covering the preceding calendar year. Very Large Online Platforms and Very Large Online Search Engines face more frequent obligations, submitting biannual reports by the end of August (covering January 1 to June 30) and by the end of February (covering July 1 to December 31).

[30] Zornetta, *supra* note 3).

[31] Ibid.

https://doi.org/10.1017/err.2026.10119

windows, allowing regulators, civil-society organisations and researchers to conduct horizontal comparisons across services and longitudinal assessments over time. Temporal harmonisation thus eliminates a structural ambiguity in the DSA's initial framework, though it remains a procedural, rather than substantive, improvement.

The second reform concerns the standardisation of the substantive categories used to classify moderation actions. In the previous iterations of transparency reports, platforms' divergent categorisation approaches prevented meaningful cross-platform comparison and verification.[32] The Implementing Regulation now mandates the use of a common and exhaustive list of high-level categories across both transparency reports and the Database. When reporting on actions based on alleged illegality with EU or national law, platforms are required to use the category of illegality indicated by competent authority.[33] Similarly, for notices reported under the notice-and-action mechanism, platforms must rely on the category corresponding to the explanation provided by the notifier.[34]

When reporting on own initiative content moderation action, platforms must use the high-level categories of illegality or the violation of the terms and conditions established in an exhaustive list.[35] Additionally, for illegal content categories, platforms have the opportunity to provide contextual information about their interpretation of the categories on the basis of their internal categorisation practices, specifying how each illegal content category description relates to the platforms' own terms and conditions policies. Platforms can also create new sub-categories to provide further context, but the Commission clarified that new sub-categories cannot be "used as a catch-all categories."[36]

By aligning the taxonomy used in transparency reports with the categories applied in the Database, the Implementing Regulation establishes for the first time a consistent vocabulary across aggregate and decision-level disclosures.[37] This structural alignment theoretically lays the groundwork for cross-mechanism verification, enabling regulators to test whether the aggregate figures reported under Articles 15 and 24 correspond to the individual moderation decisions logged in the Database.

Beyond harmonising reference periods and substantive categories, the Implementing Regulation standardises the format of transparency reports, prescribing not only what platforms must disclose but how this information must be presented.[38] Additionally, the Implementing Regulation also defines uniform data-typing conventions, limiting the potential for variations in units and denominations that made the first DSA reports virtually incomparable.[39]

[32] Ibid.

[33] DSA, Art. 15(1)(a).

[34] Ibid, Arts. 15(1)(b) and 16.

[35] Annex II, *Implementing Regulation,* pp 6–11. At the time of writing, the list contains seventeen categories: animal welfare; consumer information infringements; cyber violence; cyber violence against women; data protection and privacy violations; illegal or harmful speech; intellectual property infringements; negative effects on civil discourse or elections; protection of minors; risk for public security; scams and/or fraud; self-harm; unsafe, non-compliant or prohibited products; violence; other violation of provider's terms and conditions; type of illegal content not specified by the public authority; and type of alleged illegal content not specified by the notifier.

[36] Annex II, *Implementing Regulation,* p 5.

[37] Zornetta, *supra* note 11.

[38] *Implementing Regulation*, Art. 1(1), (requiring platforms to "use the CSV- or XLXS-version of the templates set out under Annex I to this regulation").

[39] Annex II, *Implementing Regulation*, point 4, p 3. The formatting requirements also clarify that platforms must use "integer numbers when dealing with counting indicators," that "all the indicators reporting a percentage shall be reporting as floating numbers in the [0,1] interval" and that "indicators on median time, shall be indicated in hours."

https://doi.org/10.1017/err.2026.10119

Collectively, the prescribed templates eliminate the interpretive discretion that previously allowed platforms to comply formally with Article 15 DSA while presenting incomparable data. These changes bring long-needed procedural coherence to the DSA's transparency regime, ensuring that platforms report within the same timeframes, use the same categorical vocabulary, and present information in the same technical format. Yet, while transparency reporting templates and harmonised rules enhanced consistency and comparability, platforms' disclosures remain self-generated, perpetuating the reliability problem.

2. *Database template modifications*

The second pillar of the 2025 reform concerns the Database of Statements of Reasons structure, which the Commission updated to ensure alignment with the new transparency-reporting templates entering into force on 1 July 2025. The revision of the SoR template and of the API schema governing submissions align the data categories used in the Database with those applied in transparency reports, creating a continuous reporting chain from individual decisions to aggregate statistics. In principle, this harmonisation could transform the Database from a stand-alone repository into a component of an integrated transparency framework.[40]

The SoR template follows (at most) the requirements under Article 17, translating its qualitative requirements into mostly pre-defined attributes and machine-readable fields.[41] The 2025 revision reorganises these attributes into a clearer hierarchy covering (i) information about the affected content, (ii) the type, duration and territorial scope of the restriction, (iii) the factual basis and origin of the investigation, (iv) the use of automated means, (v) the legal or contractual ground, and (vi) the categorisation of the violation.[42] Three aspects of this reorganisation are particularly consequential for assessing whether the Database can contribute to verification rather than achieve pure documentation.

First, the section on decisional grounds now requires platforms to specify, in all cases, whether a moderation action is based on illegality or incompatibility with the provider's terms and conditions. When a decision relies on a finding of illegality, platforms must indicate the precise legal provision invoked (illegal_content_legal_ground) and provide an accompanying explanation of its applicability (illegal_content_explanation).[43] When the restriction stems from a breach of contractual terms, they must cite the relevant section of the terms of service (incompatible_content_ground) and justify why the content violates it (incompatible_content_explanation).[44] These are now mandatory fields, replacing what had previously been optional or inconsistently filled. The new guidance clarifies that platforms must refer to the exact law or contractual clause relied upon and explain its relevance to the decision.[45] This formulation represents a meaningful step forward, as it formally prevents generic statements such as "incompatible with T&Cs" and compels reference to the specific provision invoked. Nevertheless, the reform continues to depend

[40] "Harmonised Transparency Reporting Rules under the Digital Services Act Now in Effect | Shaping Europe's Digital Future," *supra* note 27.

[41] R. Kaushal and others, "Automated Transparency: A Legal and Empirical Analysis of the Digital Services Act Transparency Database" *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (Association for Computing Machinery 2024) <https://doi.org/10.1145/3630106.3658960> (last accessed 9 September 2024).

[42] "API and Schema – DSA Transparency Database" <https://transparency.dsa.ec.europa.eu/page/api-documentation?lang=en#statement-attributes> (last accessed 26 October 2025).

[43] Ibid.

[44] Ibid.

[45] "Overview Documentation – DSA Transparency Database," *supra* note 18.

https://doi.org/10.1017/err.2026.10119

on self-declared accuracy. Even with more prescriptive fields, providers may comply nominally while maintaining opacity through minimal or vague legal citations ("Article 10 law," "Section 3 policy violation") that defy meaningful verification.[46] Furthermore, fields that could force further specificity – such as the reference URL allowing platforms to link directly to the cited legal or contractual provision – remain optional (decision_ground_reference_url).[47] The first analyses of the Database already highlighted the risk in leaving this field optional, with research showing that platforms rarely provide precise references for non-mandatory fields.[48]

Second, the template retains the structure governing the disclosure of the facts and circumstances relied upon in taking a moderation decision. The free-text field (decision_facts) remains mandatory, complemented by the structured indicator identifying the source of the investigation (source_type), i.e., whether the decision originated from a user notice under Article 16, a trusted flagger, another external notification, or the platform's own initiative.[49] Because the revision leaves the substance of the requirement unchanged, the limitations observed in the Database's first iteration may persist. Analyses of early entries found that, although the "decision_facts" field was mandatory and central to contextualising content moderation decisions, it was mostly filled with generic or templated responses that offered little meaningful insight.[50]

Third, the categorisation of violations has undergone the most visible restructuring. The Database now mirrors the taxonomy established in Annex II to the Implementing Regulation, requiring each statement of reasons to include one mandatory high-level category (statement_category) and, optionally, additional categories (statement_category_addition) and specifications (category_specification).[51] This exhaustive, harmonised list spans seventeen high-level categories and over eighty subcategories. Providers may still create new subcategories under "keyword_other," but must supply a textual description of the content type and cannot add new high-level categories.[52] This system is intended to prevent the proliferation of idiosyncratic classifications that, as previously documented, hindered cross-platform comparability and frustrated the objective of "meaningful transparency" under Article 15 DSA.[53] However, this controlled flexibility does not prevent the risk for abuse. Over-reliance on open subcategories could reintroduce fragmentation by allowing providers to label edge cases under self-defined headings ("other misinformation," "borderline nudity") that escape harmonised scrutiny.[54]

[46] Zornetta, *supra* note 3, 42. Under the original Database schema (Feb 2024), these explanatory fields were optional, allowing platforms to state merely "violates applicable law" or "breach of terms of service." The 2025 revision closes that loophole by making the fields mandatory, but without validation of their substance, platforms can still comply superficially, for example by citing only "EU hate speech laws" or "our harmful content policy."

[47] API Guidance: (decision_ground_reference_url) is optional.

[48] Kaushal and others, *supra* note 41, finding that fewer than 0.1 percent of entries included a reference URL.

[49] "Overview Documentation – DSA Transparency Database," *supra* note 18; "API and Schema – DSA Transparency Database," *supra* note 42.

[50] Kaushal and others, *supra* note 41, 8. ("Google Shopping, the most active VLOPs, submits only a standard template8 in most (99.5%) of the SoRs. The next key contributor of SoRs, TikTok, uses a similar albeit shorter statement [The decision was taken pursuant to own-initiative investigations.'] in 98.5% of its SoRs.").

[51] "Implementing Regulation," *supra* note 12.

[52] Ibid.

[53] Zornetta, *supra* note 11; Trujillo, Fagni and Cresci, *supra* note 25; Kaushal and others, *supra* note 41.

[54] Even though the Implementing Regulation explicitly warns providers not to use "keyword_other" as a catch-all category, compliance with this restriction will depend on the Commission's monitoring capacity. See in general, discussing the risks of procedural but not substantial standardisation: D Keller and P Leerssen, "Facts and Where to Find Them: Empirical Research on Internet Platforms and Content Moderation", *Social media and democracy: The state of the field and prospects for reform*, vol 220 (Cambridge University Press Cambridge, UK 2020) 4,15. Zornetta, *supra* note 11, 122.

https://doi.org/10.1017/err.2026.10119

Despite these improvements, several omissions identified in earlier analyses of the Database structure remain unresolved. The submission template still lacks fields corresponding to the redress mechanisms available to affected users, even though Article 17(3)(f) explicitly requires that statements of reasons include "clear and user-friendly information on the possibilities for redress available." Without this element, however, the Database captures only the moment of restriction, rather than the broader procedural relationship between the platform and the user. It therefore provides no means to assess whether users are informed of their right to contest a decision, how platforms operationalise internal complaint mechanisms, or whether redress avenues are offered consistently across jurisdictions and content types. Likewise, the Database still provides no field to record whether a decision was later challenged through an internal complaint or out-of-court settlement, nor any mechanism to update or withdraw an SoR once a decision has been reversed.[55] Consequently, what is recorded in the Database is a static snapshot of initial decisions rather than the full lifecycle of a moderation decision, providing an incomplete and potentially misleading picture of platform practices.

Overall, the 2025 revisions improve the internal coherence of the Database template and strengthen its alignment with the transparency reporting framework. Yet these gains remain incremental rather than transformative. The Database continues to serve primarily as a documentation tool rather than a verification mechanism. The reforms ensure that platforms describe their decisions in a uniform manner, but they do not ensure that those descriptions are complete or accurate. In other words, harmonisation mitigates inconsistency but not unreliability. The Database now speaks the same language as transparency reports, yet it still cannot independently confirm whether what platforms say is true.

IV. The database's untapped potential as a verification tool

The 2025 Implementing Regulation and the accompanying updates to the Statement of Reasons (SoR) template represent a substantial advancement toward a more harmonised transparency regime. Despite these improvements, the DSA's transparency framework remains essentially descriptive. The Database continues to operate as a repository of disclosures designed for observation, not verification, leaving the underlying reliability problem unsolved. As in the first generation of DSA transparency reports, accuracy remains presumed rather than demonstrated. This limitation is fundamental. The Database now records billions of individual content moderation decisions, but these granular records remain disconnected from the aggregated data presented in transparency reports. Without a formalised mechanism to reconcile the two, regulators cannot determine whether the figures disclosed in transparency reports accurately reflect the respective moderation activity.

This article argues that, if properly leveraged, the Database could provide the systematic verification necessary for meaningful transparency, by allowing regulators to cross-check reported aggregated figures against individual SoR submissions, detect inconsistencies, and prioritise investigative resources. This would not only enhance accountability for VLOPs and VLOSEs, already subject to audits, but – perhaps more importantly – extend verifiability to the broader ecosystem of intermediaries that fall outside the auditing framework but must still comply with the same transparency obligations.

This section explains how the Database's current underutilisation stems not from its technical incapacity but from the absence of regulatory design for verification. It proceeds in three steps. Section IV.1 examines how the Database's architecture and the recent

[55] Ibid, 119–25.

https://doi.org/10.1017/err.2026.10119

template updates have improved its harmonisation but left structural disconnections unaddressed. Section IV.2 discusses why the verification function is particularly crucial for smaller intermediaries and Digital Services Coordinators (DSCs). Section IV.3 identifies the legal, technical, and institutional preconditions required for the Database to evolve into a verification tool capable of reinforcing the reliability of the DSA transparency framework as a whole.

1. From documentation to verification

The Implementing Regulation marks a turning point in the DSA's transparency framework. For the first time, transparency reports and statements of reasons share compatible, machine-readable structures. Yet a significant gap remains: there is no mechanism linking individual SoRs to the aggregated data in transparency reports, nor any means of automatically testing the consistency between the two datasets.

The Implementing Regulation nevertheless brings the system within reach of verifiability. First, the insistence on machine-readability creates a data architecture that can support automated cross-analysis. Second, Annex II's exhaustive taxonomy of illegal-content categories (seventeen high-level types, each subdivided into specific sub-categories) aligns the categorisation logic of transparency reports with that of the Database. Building on this comparability structure, a limited set of additional obligations would suffice to transform the Database from a passive documentation tool into an active oversight mechanism. These obligations would focus on introducing traceability across the two transparency mechanisms: (i) mandatory cross-referencing fields linking aggregated figures to Database submissions; (ii) update fields in SoRs for corrected or appealed decisions; and (iii) automatic aggregation checks at the time of report submission. Absent such requirements, however, the Database will remain a purely descriptive exercise rather than a verifiable transparency system.

2. Why verification matters beyond VLOPs/VLOSEs

The DSA's auditing obligations apply only to VLOPs and VLOSEs. These entities are subject to independent annual audits under Article 37, which evaluate their compliance with risk-assessment and mitigation duties, including the accuracy of their transparency reporting. By contrast, all other intermediaries remain outside the auditing framework. They are nonetheless bound by the same Article 15 transparency obligations and therefore continue to self-report moderation data without external verification.

Integrating verification functions within the Database would help fill this gap. Because all intermediaries are required to submit SoRs, the Database already collects granular data capable of supporting basic reliability checks. Automated consistency analyses – performed either by the Commission or by Digital Services Coordinators (DSCs) – could compare the volume, distribution, and categorisation of individual decisions against the aggregated statistics reported under Articles 15 and 24. Even simple anomaly-detection tools could flag inaccurate patterns, such as missing categories, inaccurate totals, or inconsistent breakdowns across reporting mechanisms.[56]

Embedding automated consistency checks within the Database infrastructure could reveal discrepancies at the time of submission, allowing DSCs to flag inconsistencies and request further information. Such checks would not replicate the depth of Article 37 audits but would create a proportional verification layer for non-VLOPs. They would enable regulators to prioritise oversight where discrepancies are most pronounced and, over

[56] Ibid, 143. Showing that entire content categories reported in transparency reports are sometimes absent from Database submissions and vice-versa. For instance, while every platform analysed in the article included "Mis-/Disinformation" among its reported moderation categories, none categorised any Database entries as such.

https://doi.org/10.1017/err.2026.10119

time, incentivise systemic accountability across the intermediary ecosystem. Database-based verification would extend the logic of auditability to all providers, transforming formal compliance into meaningful transparency.

3. Structural preconditions for verification

Transforming the Database into a functional verification instrument would require three interdependent preconditions: technical, legal and institutional. For what concerns the technical foundation, the harmonisation process has already created the structural compatibility necessary for automated reconciliation between transparency reports and Database submissions. This harmonisation ensures that both instruments are expressed in compatible, machine-readable structures and that their variables can be compared automatically. If regulators were to analyse these aligned datasets in parallel, they could – theoretically – detect inconsistencies between what platforms report in aggregate and what they record at the decision level. The technical precondition for verification is therefore already in place but without a traceability requirement, these features will remain underutilised in oversight.

On the other hand, the legal precondition for verification is absent from the DSA transparency framework. Neither the DSA nor the Implementing Regulation explicitly require providers to ensure internal consistency between their transparency reports and the SoRs transmitted to the Database. Article 17 imposes a duty to communicate explanations for individual moderation decisions, while Articles 15 and 24 require the publication of aggregate transparency data, but the two obligations operate independently. Similarly, the Implementing Regulation focuses on harmonising format and taxonomy, not on establishing mandatory coherence or cross-verification. As a result, even though the datasets are now structurally comparable, platforms are under no legal obligation to demonstrate that their reported figures accurately reflect the decisions recorded in the Database, and regulators lack a clear legal mandate to test such consistency. Introducing verification would therefore require an interpretative clarification or secondary measure establishing a traceability requirement, i.e., the expectation that aggregated data must be reconcilable with the underlying statements of reasons. Without such an interpretative step, verification remains permissible in theory but not enforceable in practice.

Finally, the verification process would require institutional capacity. The DSA already distributes oversight responsibilities across the Commission and Digital Services Coordinators, empowering both with investigative and information-request powers.[57] On the other hand, there is no structured process to prioritise their use of those powers. For VLOPs and VLOSEs, independent audits provide an initial verification cycle. On the contrary, for smaller intermediaries, oversight remains reactive, triggered only by complaints or ad hoc suspicion of non-compliance. Establishing an intermediate verification process could make this oversight more systematic. Ideally, when platforms submit their annual transparency reports, the figures disclosed would be automatically matched to the statements of reasons they uploaded to the Database during the same reporting period.[58] This could occur through an interface integrated with the Database, allowing DSCs to view a high-level consistency summary for each provider. Platforms showing unexplained discrepancies would be flagged for further review.

[57] See *supra* Section II.

[58] Two studies have already attempted such reconciliation exercises before the adoption of the harmonised templates. Trujillo, Fagni and Cresci, *supra* note 25, examining the Database entries of the eight largest platforms during its first 100 days and finding data-quality issues and structural inconsistencies, notably X's reporting discrepancies); and Zornetta, "The DSA Database as a Tool for Transparency Accountability," *supra* note 11 (conducting a comparative analysis of Database entries and respective transparency reports from seven VLOPs).

https://doi.org/10.1017/err.2026.10119

Such a system would not create new investigatory powers but would help regulators use existing ones more strategically, enabling proactive oversight and evidence-based prioritisation of enforcement efforts.

4. Mandating traceability

The preceding analysis indicates that the harmonisation process already laid the foundation for the verification of accuracy of transparency reports through cross-comparison with Database entries. To operationalise the structural preconditions identified in Section IV.3, a legal basis linking the different transparency instruments of the DSA into a coherent oversight framework is needed. Addressing this gap does not necessarily require a redesign of the Database or the creation of a new reporting layer. Rather, it calls for the introduction of an explicit traceability requirement establishing that the aggregated information disclosed under Articles 15 and 24 must be reconcilable with the statements of reasons transmitted under Article 17.

This requirement could be articulated through a delegated act under Article 15 or Article 24, both of which empower the Commission to specify the form, content and details of reporting obligations. Such an instrument could clarify that traceability forms part of the duty of accuracy inherent in transparency reporting, thereby allowing verification to occur within the existing legal framework.

A traceability requirement would not expand the substantive scope of the DSA's reporting duties. Providers would not be required to submit additional information but to ensure that the data they already transmit remain internally coherent across mechanisms. The obligation could be framed procedurally: when publishing transparency reports, providers must be able, upon request by DSCs or the Commission, to demonstrate how the reported figures align with the SoRs submitted during the relevant reporting period. This clarification would give legal effect to the semantic alignment introduced by the Implementing Regulation, transforming it from a technical feature into a compliance benchmark. It would also provide a defined legal basis for regulators to perform consistency analyses and to request methodological explanations where anomalies arise.

Formally articulating the relationship between Articles 15, 17 and 24 would also support a more coherent use of existing investigatory powers. Under Articles 51–53, both the Commission and DSCs may already request information and conduct inspections, but these powers operate reactively. A traceability requirement would create an intermediary verification stage, allowing DSCs to use the Database as a first-line screening tool before initiating formal proceedings. Platforms whose reported data diverge significantly from their Database entries could be prioritised for investigation or asked to provide clarifications under Article 52(1). This would strengthen proportionality and resource allocation in enforcement while maintaining the DSA's graduated oversight structure.

V. Conclusion

The July 2025 reforms mark a significant step toward a coherent transparency framework under the DSA. By harmonising temporal scopes, taxonomies and reporting formats, the Implementing Regulation and the corresponding updates to the Database template address many of the inconsistencies that undermined the first transparency cycle. These measures enhance comparability and coherence across instruments, creating the structural preconditions for a verifiable transparency regime. Yet, the reliability problem – the persistent inability to verify whether the information platforms disclose accurately reflects their moderation practices – remains the central weakness of the DSA's transparency regime. The DSA already contains the tools and structure necessary to

https://doi.org/10.1017/err.2026.10119

address this issue: the Database of Statements of Reasons provides granular, decision-level data that could enable systematic verification of reported figures. However, the 2025 reforms stop short of operationalising this potential. The Database remains a powerful but underutilised instrument, technically capable of supporting verification but limited to an observation tool.

Closing this gap would not require a new reporting layer or a fundamental redesign of the existing mechanisms. A narrowly tailored duty of traceability could ensure that the data reported under Articles 15 and 24 are reconcilable with the statements of reasons transmitted under Article 17. Such a clarification would translate the harmonisation achieved through the 2025 reforms into operational verifiability, allowing regulators to assess whether the reported aggregates correspond to the underlying decision-level data. Traceability would not expand the scope of reporting but would give legal effect to the expectation of internal coherence already implicit in the DSA's transparency obligations.

Embedding traceability would enhance both proportionality and oversight. For VLOPs and VLOSEs, it would complement existing auditing duties by facilitating independent cross-checks; for smaller intermediaries, it would introduce a baseline verification mechanism within the same legal framework. In this way, traceability would extend accountability beyond the audited entities, turning transparency into a reliable measure of platform conduct rather than a self-declared account of it. The DSA's transparency architecture would thus evolve from one that records platform behaviour to one that can substantively verify it, realising its original purpose as an instrument of trustworthy and enforceable accountability.

Dr Alessia Zornetta holds an SJD from the University of California, Los Angeles, an LLM from McGill University, and an LLB from the University of Trento. She is affiliated with the UCLA Institute for Technology, Law & Policy.
The author wishes to thank her SJD supervisor, Prof. John Villasenor, as well as Prof. Michael Karanicolas, Prof. Sarah T. Roberts, and Prof. Yuan Tian for their invaluable advice and support.

Cite this article: A Zornetta, "Turning the DSA into a Verification Tool: A Feasible Solution to the Reliability Problem". *European Journal of Risk Regulation*. https://doi.org/10.1017/err.2026.10119

https://doi.org/10.1017/err.2026.10119