I6EAL RESEARCH NOTES IN AI ECONOMICS AND DIGITAL POLICY

# Automated Detection and Automated Decision-Making Under the EU Digital Services Act: Evidence from 43.5 Billion Statements of Reasons (2023–2026)

Ideal Syka · i6eal / Syka Ventures UG (haftungsbeschränkt)

i6eal-RN-2026-02 · 18 November 2026 · English · CC BY 4.0

ISSN 3056-0489 · VERSION 1.0

| 43.46B | 93.86% | 54.50% | +39.35 pp |
|---|---|---|---|
| published SoRs | automated detection | fully automated | detection–full gap |

## ABSTRACT

This research note examines the empirical relationship between reported automated detection and reported automated decision-making across the first 38 monthly aggregate releases of the European Union Digital Services Act (DSA) Transparency Database (25 September 2023 to 31 October 2026; 43,463,346,962 published Statements of Reasons across 278 named platforms). Across the cumulative aggregate, 93.86% of statements report automated detection, whereas 54.50% report fully automated decision-making, 9.60% partially automated decision-making and 35.90% non-automated decision-making—a 39.35 percentage-point gap between automated intake flagging and full decision automation. In the joint 2x3 contingency table, 34.50% of automatically detected statements (14.07 billion records) are recorded as non-automated at the decision stage. Provider-level decomposition shows that unweighted EU totals are heavily shaped by extreme concentration: Google Shopping accounts for 76.95% of all published statements and 94.96% of all non-automated decision records. Across the remaining 277 platforms (10.02 billion statements), a 34.65 percentage-point gap persists between automated detection (85.14%) and fully automated decisions (50.49%), while partially automated decisions account for 41.66% and non-automated decisions for 7.85%. All results are descriptive accounting identities over provider-submitted administrative records; statutory 'automated means' does not establish artificial-intelligence use, review quality or DSA compliance.

JEL L86, K24, O33, D83
KEYWORDS Digital Services Act · content moderation · platform governance · automated decision-making · automated detection · transparency database · European Union

# 1. Research question and contribution

Under Article 17 of the European Union Digital Services Act (Regulation (EU) 2022/2065, DSA), providers of hosting services and online platforms submit machine-readable Statements of Reasons whenever they restrict the visibility, monetisation, service access or account status of user-supplied information. Under Article 24(5) DSA, the European Commission publishes those records in the DSA Transparency Database. Unlike earlier voluntary platform transparency reports, the database records two distinct automation attributes on every submission: whether the content was detected or identified using automated means, and whether the subsequent moderation decision was taken using fully automated, partially automated or non-automated means.

Public commentary and aggregate policy summaries frequently collapse those two stages into a single headline share of automated moderation, or treat aggregate EU database totals as representative of typical platform practice. Both shortcuts obscure the structure of the data. First, automated flagging at intake does not imply unreviewed automated execution at the decision stage. Second, statement volume in the DSA Transparency Database is extraordinarily concentrated, so a single high-volume commerce service can dominate unweighted European percentages.

This note asks three narrow descriptive questions across the first 38 monthly aggregate releases of the DSA Transparency Database (25 September 2023 to 31 October 2026; 43,463,346,962 published Statements of Reasons across 278 named platforms): (1) how large is the separation between reported automated detection and reported automated decision-making in the cumulative aggregate and in the joint 2x3 detection-by-decision contingency table; (2) how much of the aggregate decision distribution reflects extreme provider concentration rather than cross-platform uniformity; and (3) how do major platforms distribute records across fully automated, partially automated and non-automated decision categories? All findings are descriptive accounting identities over provider-submitted administrative records and make no causal, welfare, accuracy or compliance claims.

**Headline result**

**Across 43,463,346,962 published Statements of Reasons (September 2023 to October 2026), 93.86% report automated detection, whereas 54.50% report fully automated decision-making, 9.60% partially automated decision-making and 35.90% non-automated decision-making: a 39.35 percentage-point gap between automated detection and full decision automation. Excluding Google Shopping (76.95% of all records), the remaining 277 platforms (10.02 billion statements) still exhibit a 34.65 percentage-point gap (85.14% automated detection versus 50.49% fully automated decisions), but shift overwhelmingly toward partially automated decisions (41.66%) rather than non-automated decisions (7.85%).**

# 2. Data, institutional scope and definitions

The analysis uses the official simplified monthly aggregate archives (aggregated-simple) published by the European Commission for the DSA Transparency Database, as compiled and checksum-verified by the i6eal EU Platform Automation Monitor (archive revision 88, covering record creation dates from 25 September 2023 through 31 October 2026). Through May 2026, the Commission distributed monthly archives as Spark-generated CSV bundles; from June 2026 onward, the official distribution switched to Apache Parquet. Across all 38 monthly archives, 2,439,993 grouped source rows sum to exactly 43,463,346,962 published Statements of Reasons across 278 platform identifiers.

The unit of observation is a published Statement-of-Reasons database record timed by its database creation month (created_at). One statement record is not an independently deduplicated piece of content, user account or appeal outcome, and multiple restriction types can co-occur within a single

submission. Crucially, the statutory phrase 'automated means' in the DSA reporting schema covers any algorithmic, rule-based, hash-matching, keyword or statistical tool used in detection or decision routing; neither automated_detection nor automated_decision establishes the deployment of machine-learning or generative artificial-intelligence systems.

- Observation window: 38 monthly periods from 2023-09 (partial launch month: 25–30 September 2023) to 2026-10 (31 October 2026).
- Detection attribute (automated_detection): binary classification into automated detection ('Yes': 40,793,188,734 statements) and non-automated detection ('No': 2,670,158,228 statements), with zero unclassified records in the frozen aggregate.
- Decision attribute (automated_decision): three-way classification into 'Fully automated' (23,688,393,399), 'Partially automated' (4,173,540,726) and 'Not automated' (15,601,412,837).
- Platform harmonisation: one deterministic source-escaping normalisation joins Spark-escaped CSV quotes ('SIA \JOOM\""' in May 2024–May 2026) with the identical Parquet string ('SIA "JOOM"' from June 2026 onward); no fuzzy name matching is used.

**Comparability and institutional timeline boundaries**

**Four documented institutional changes shape longitudinal totals: (1) on 17 February 2024, reporting obligations expanded from designated Very Large Online Platforms (VLOPs) to other non-exempt online platforms; (2) Google Shopping narrowed its notification scope on 28 June 2024 (personalised-ad eligibility) and 30 May 2025 (additional grounds), reducing aggregate monthly volume; (3) Schema v2 launched on 1 July 2025 without altering the automation-field definitions; and (4) on 29 May 2026 (Commission notice 8 June 2026), Google corrected a historical inversion of automated-versus-manual detection flags on a subset of Google Play submissions from September 2023 to May 2026.**

# 3. Method

Because the source data represent the complete population of published aggregate rows rather than a probability sample, the method consists of exact accounting decompositions across three views of the frozen projection: the marginal and joint 2x3 contingency table of automated_detection by automated_decision (summed across 228 monthly flow cells), a concentration decomposition separating the largest reporting service (Google Shopping) from the remaining 277 platforms, and a cross-platform comparison of the ten largest services by cumulative statement volume.

For each slice, percentage shares use exact integer statement counts in the numerator and denominator before rounding to two decimal places. We define the primary detection-to-full-decision gap as the percentage of statements reporting automated detection minus the percentage reporting fully automated decision-making (in percentage points, pp), alongside a secondary gap relative to any automated decision involvement (fully plus partially automated). No reweighting, smoothing or statistical inference is applied.

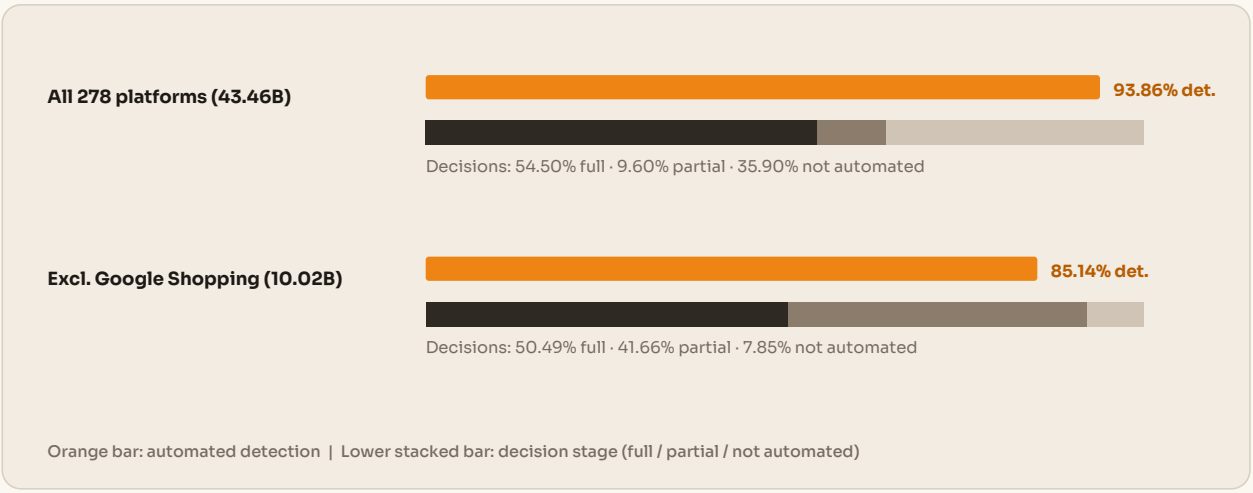



**Figure 1.** Automated detection share versus three-way decision-stage distribution across all 278 platforms (43.46B statements) and across the 277 platforms excluding Google Shopping (10.02B statements), 2023-09 to 2026-10.
Source: European Commission DSA Transparency Database (aggregated-simple, rev. 88); frozen i6eal analysis table.

# 4. Results

## 4.1 Joint 2x3 contingency structure of detection and decision-making

At the European aggregate level, 40,793,188,734 of 43,463,346,962 statements (93.86%) report automated detection, while 2,670,158,228 (6.14%) report non-automated detection. Decision-stage declarations are substantially less concentrated in full automation: 23,688,393,399 statements (54.50%) report fully automated decisions, 4,173,540,726 (9.60%) report partially automated decisions and 15,601,412,837 (35.90%) report non-automated decisions. Combining fully and partially automated decisions yields 64.10% with at least partial decision automation, still 29.75 percentage points below the automated-detection share.

Cross-tabulating both attributes simultaneously clarifies how intake and decision workflows connect in provider submissions. Conditional on automated detection (40.79 billion records), 57.75% (23,557,915,394; 54.20% of all database records) proceed to a fully automated decision, 7.75% (3,163,124,037; 7.28% of total) proceed to a partially automated decision, and 34.50% (14,072,149,303; 32.38% of total) are recorded as non-automated decisions. Conversely, conditional on non-automated detection (2.67 billion records), 57.27% (1,529,263,534) remain non-automated at the decision stage, 37.84% (1,010,416,689) involve partially automated decisions, and only 4.89% (130,478,005; 0.30% of total) report a fully automated decision.

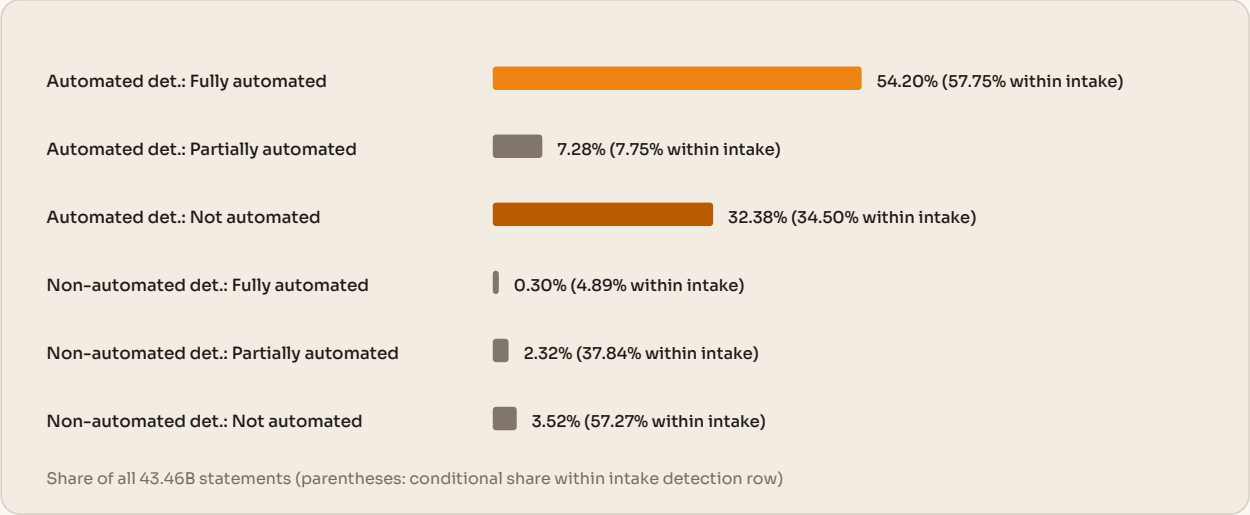



**Figure 2.** Joint 2x3 distribution of automated_detection and automated_decision across 43,463,346,962 published Statements of Reasons.
Source: European Commission DSA Transparency Database (aggregated-simple, rev. 88); frozen i6eal analysis table.

## 4.2 Provider concentration and the Google Shopping composition effect

Unweighted aggregate percentages in the DSA Transparency Database are dominated by extreme skewness across services. The single largest service, Google Shopping, accounts for 33,445,725,856 statements—76.95% of the entire 38-month corpus. The top five services (Google Shopping, TikTok, Amazon, Facebook and AliExpress) account for 92.50%, and the top ten services account for 97.29% (42,283,321,593 statements) across 278 reporting platforms.

Separating Google Shopping from the rest of the database materially alters how the aggregate non-automated decision share must be interpreted. Google Shopping reports 96.47% automated detection, 55.70% fully automated decisions, 0.00% partially automated decisions and 44.30% non-automated decisions (14,815,969,624 statements). Consequently, Google Shopping alone generates 94.96% of all 15.60 billion non-automated decision records in the European Union aggregate. Across the remaining 277 platforms combined (10,017,621,106 statements; 23.05% of total), automated detection is 85.14% and fully automated decision-making is 50.49%—preserving a large 34.65 percentage-point gap—but the non-full remainder is recorded predominantly as partially automated decision-making (41.66%) rather than non-automated decision-making (7.85%).

## 4.3 Calendar-period evolution and platform reporting regimes

Across calendar periods, total volume peaked in 2024 (22.24 billion statements across up to 65 monthly platforms), declined to 12.80 billion in 2025 (up to 158 platforms) following Google Shopping's scope changes, and reached 7.02 billion across January–October 2026 (up to 213 platforms). Over those four windows, the aggregate gap between automated detection and fully automated decision-making widened from 25.98 pp in late 2023 (93.33% versus 67.35%) to 37.34 pp in 2024, 42.01 pp in 2025 and 43.56 pp in January–October 2026 (86.55% versus 42.99%), as partially automated decisions rose to 22.98% in 2026.

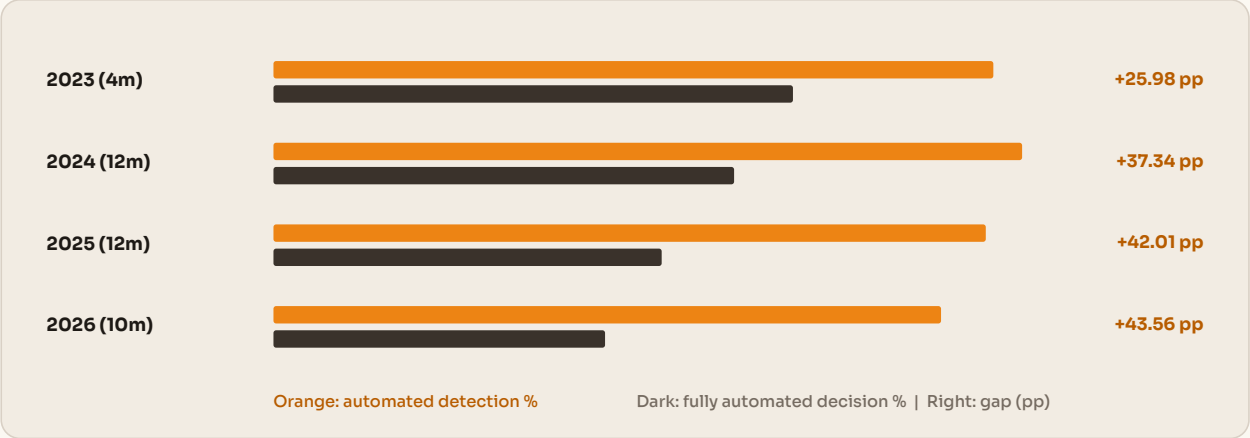



**Figure 3.** Calendar-period evolution of automated detection and fully automated decision shares, 2023 (Sep–Dec) through 2026 (Jan–Oct).
Source: European Commission DSA Transparency Database (aggregated-simple, rev. 88); frozen i6eal analysis table.

| Slice / calendar window | Statements | Auto det. | Full dec. | Partial | Not auto | Gap |
|---|---|---|---|---|---|---|
| All 278 platforms (2023-09..2026-10) | 43,463,346,962 | 93.86% | 54.50% | 9.60% | 35.90% | +39.35 pp |
| Excl. Google Shopping (277 platforms) | 10,017,621,106 | 85.14% | 50.49% | 41.66% | 7.85% | +34.65 pp |
| 2023 (2023-09 to 2023-12, peak 16 plat.) | 1,404,823,561 | 93.33% | 67.35% | 15.79% | 16.86% | +25.98 pp |
| 2024 (2024-01 to 2024-12, peak 65 plat.) | 22,240,512,622 | 97.06% | 59.72% | 5.39% | 34.89% | +37.34 pp |
| 2025 (2025-01 to 2025-12, peak 158 plat.) | 12,800,939,222 | 92.35% | 50.33% | 8.92% | 40.75% | +42.01 pp |
| 2026 (2026-01 to 2026-10, peak 213 plat.) | 7,017,071,557 | 86.55% | 42.99% | 22.98% | 34.03% | +43.56 pp |

**Table 1.** Aggregate, concentration-adjusted and calendar-period automation shares, September 2023 to October 2026. Shares are percentages of published Statements of Reasons in each slice; Gap = automated detection % minus fully automated decision % (pp).

Inspecting the ten largest platforms individually reveals four distinct reporting regimes rather than a continuum around the EU mean. First, TikTok (98.07% detection; 94.09% full), Google Maps (96.81%; 94.71%), AliExpress (89.66%; 86.78%) and Temu (99.70%; 85.40%) report near-complete alignment between automated detection and fully automated decision-making (gaps of 2.10 to 14.30 pp). Second, Meta's Facebook (97.58% automated detection) and Instagram (95.54%) report 0.00% fully automated decisions, classifying virtually all automatically detected restrictions as partially automated (97.58% and 95.54%, respectively); Amazon similarly reports 91.87% automated detection against 11.75% fully automated and 88.24% partially automated decisions. Third, Google Shopping (55.70% full; 44.30% not automated) and YouTube (44.04% full; 4.54% partial; 51.42% not automated) exhibit bimodal splits between fully automated and non-automated decisions. Fourth, Pinterest reports only 3.60% automated detection alongside 95.99% partially automated decisions, illustrating substantial cross-provider divergence in how hybrid enforcement pipelines are mapped onto the statutory schema.

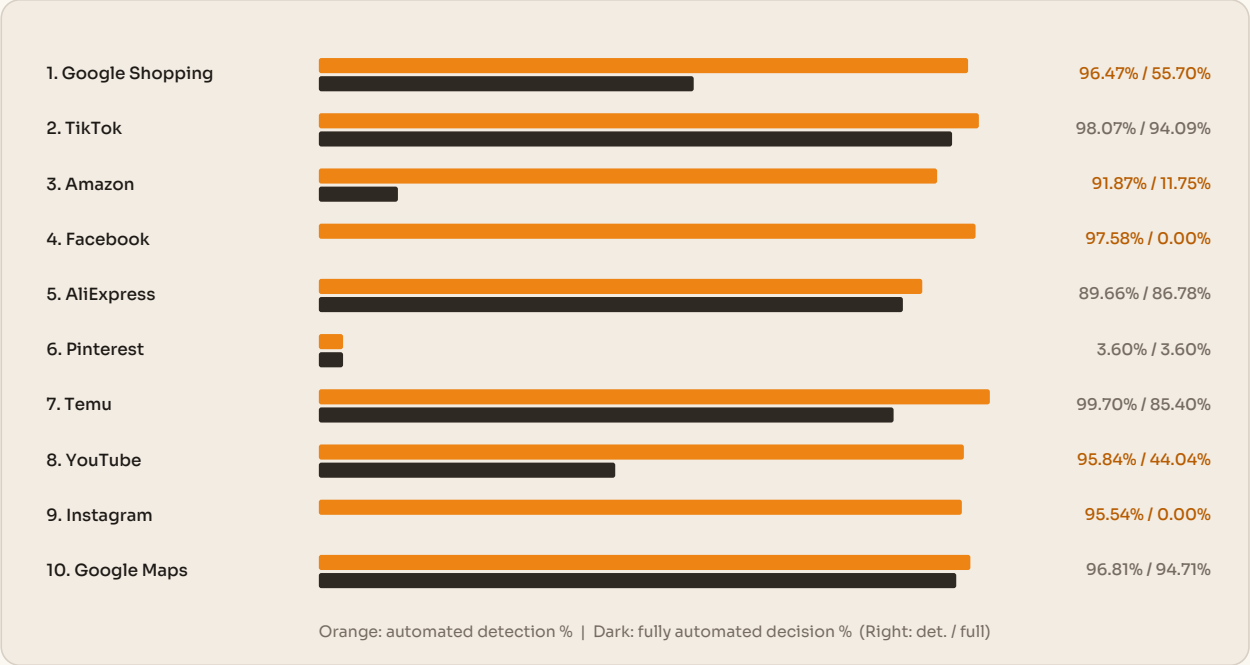



**Figure 4.** Top ten platforms by cumulative Statement-of-Reasons volume (97.29% of EU total): automated detection % versus fully automated decision %.
Source: European Commission DSA Transparency Database (aggregated-simple, rev. 88); frozen i6eal analysis table.

| Platform | Statements | Share | Auto det. | Full dec. | Partial | Not auto |
|---|---|---|---|---|---|---|
| 1. Google Shopping | 33,445,725,856 | 76.95% | 96.47% | 55.70% | 0.00% | 44.30% |
| 2. TikTok | 2,683,184,920 | 6.17% | 98.07% | 94.09% | 0.00% | 5.91% |
| 3. Amazon | 1,520,407,257 | 3.50% | 91.87% | 11.75% | 88.24% | 0.00% |
| 4. Facebook | 1,367,231,804 | 3.15% | 97.58% | 0.00% | 97.58% | 2.42% |
| 5. AliExpress | 1,186,189,756 | 2.73% | 89.66% | 86.78% | 9.29% | 3.93% |
| 6. Pinterest | 904,370,023 | 2.08% | 3.60% | 3.60% | 95.99% | 0.42% |
| 7. Temu | 361,853,424 | 0.83% | 99.70% | 85.40% | 0.23% | 14.37% |
| 8. YouTube | 296,281,958 | 0.68% | 95.84% | 44.04% | 4.54% | 51.42% |
| 9. Instagram | 284,096,935 | 0.65% | 95.54% | 0.00% | 95.54% | 4.46% |
| 10. Google Maps | 233,979,660 | 0.54% | 96.81% | 94.71% | 1.66% | 3.62% |

**Table 2.** Ten largest platforms by cumulative Statement-of-Reasons volume (42.28B statements; 97.29% of EU total).
Shares are exact statement counts divided by platform total before rounding to two decimals.

# 5. Policy and methodological interpretation

Three lessons follow for digital-market research and regulatory oversight. First, detection automation and decision automation are empirically distinct margins in DSA filings: treating a 93.86% automated-detection rate as evidence that 94% of content restrictions are executed without human involvement overstates full decision automation by nearly forty percentage points. Second, unweighted database-wide percentages should never be cited without provider-concentration controls. Because one retail-listing service accounts for more than three-quarters of all records and 95% of all non-automated decisions, shifts in a single company's reporting perimeter move the European aggregate more than the entry of two hundred smaller platforms.

Third, the stark polarization among major platforms—such as 0.00% fully automated decisions at Facebook and Instagram versus 94.09% at TikTok, or 95.99% partially automated decisions despite 3.60% automated detection at Pinterest—suggests that institutional definitions of 'partially automated' and 'not automated' carry substantial interpretive latitude across compliance teams. Cross-platform comparisons of human-in-the-loop oversight therefore require harmonised audit guidance rather than literal face-value comparisons of raw category shares.

# 6. Limitations

- Statements of Reasons are self-reported by platform providers; neither the European Commission's simplified aggregate nor this study independently audits whether reported human involvement was substantive or nominal.
- Statutory 'automated means' encompasses deterministic hashes, keyword rules and metadata heuristics as well as machine-learning models; these figures do not measure AI adoption.
- Counts reflect database submissions by creation month (created_at), not unique content items, unique users, underlying infraction dates or post-appeal reversals.
- The simplified aggregate excludes territorial breakdowns, content-category cross-tabs within flow cells and statement-level standard errors.
- Longitudinal series reflect changing reporting populations (February 2024 VLOP expansion), provider scope changes (Google Shopping in June 2024 and May 2025), Schema v2 migration (July 2025) and historical provider bugs (Google Play detection inversion prior to 29 May 2026).

# 7. Conclusion

Across 43.46 billion Statements of Reasons published during the first 38 months of the EU DSA Transparency Database, automated detection is near-universal (93.86%), whereas fully automated decision-making accounts for just over half of submissions (54.50%). Decomposing the 39.35 percentage-point gap shows both a genuine two-stage workflow separation across platforms and an extraordinary composition effect: Google Shopping represents 76.95% of all submissions and 94.96% of all non-automated decision records, while the remaining 277 platforms split automated intake primarily between fully automated (50.49%) and partially automated (41.66%) decisions.

Future empirical work on platform governance under the DSA should report detection and decision stages jointly, always present concentration-adjusted or platform-disaggregated figures alongside database totals, and treat cross-provider differences between 'partially' and 'fully' automated decisions as a combination of operational design and reporting-schema interpretation.

# Data availability and reproducibility

The frozen analysis CSV, bibliographic metadata, BibTeX and RIS citations, and SHA-256-fingerprinted source tables (archive revision 88) accompany this paper at https://i6eal.de/research-notes/automated-detection-decision-making-eu-dsa/. Ongoing monthly updates are tracked by the i6eal EU Platform Automation Monitor at https://i6eal.de/tools/plattformautomationsmonitor/ and https://i6eal.de/data/plattformautomationsmonitor/methodology.json. Upstream source aggregates are published by the European Commission under CC BY 4.0; this paper and the i6eal analytical compilation are licensed under CC BY 4.0.

# References

— European Parliament and Council of the European Union (2022). Regulation (EU) 2022/2065 of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act), Articles 17 and 24(5). Official Journal of the European Union, L 277, 1–102. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=oj:JOL_2022_277_R

— European Commission (2023). Digital Services Act: Commission launches Transparency Database. 25 September 2023. https://digital-strategy.ec.europa.eu/en/news/digital-services-act-commission-launches-transparency-database

— European Commission (2025). DSA Transparency Database — Statement attribute definitions and API changelog (Schema v1 to v2). https://transparency.dsa.ec.europa.eu/page/additional-explanation-for-statement-attributes?lang=en

— European Commission (2026). DSA Transparency Database — Announcements (including Google Shopping scope notifications and the 8 June 2026 Google Play detection-flag notice). https://transparency.dsa.ec.europa.eu/page/announcements?lang=en

— European Commission (2026). DSA Transparency Database — Download and simplified aggregate documentation (dsa-tdb). Accessed 18 November 2026. https://dsa.pages.code.europa.eu/transparency-database/dsa-tdb/data_sources.html