

Beyond Platform Transparency: Ten Policy Recommendations for Digital Public-Interest Research

DIGIPOL – Deliverable D10

Zach Bastick
European University Institute (EUI), Florence
02 June 2025

Funding statement

Funded by the European Union. Views and opinions expressed are, however, those of the author only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the granting authority can be held responsible for them.



Funded by the
European Union

Foreword

This policy paper disseminates lessons learned from DIGIPOL, the Digital Observatory of Political Research project funded by the European Commission. The project began from a premise that was, at its start, methodologically familiar and politically necessary: that it would be possible to build a data-driven observatory of platform dynamics by relying on platform-accessible interfaces, initially centered on Twitter. The project's trajectory, however, exposed a more fundamental governance problem. As access conditions shifted across social media platforms through API restrictions, changing eligibility rules, pricing decisions, interface redesign, and the deprecation of third-party monitoring tools, the practical basis for independent and continuous observation became unstable. This emphasized our focus on the governance of platform access – an emergent issue of the present era of digital research, in which public interest research can no longer rely on an open-door policy.

DIGIPOL therefore offers a lesson that extends well beyond a single project or platform. The central issue is not merely that researchers sometimes lose convenient tools. Rather, it is that democratic oversight of core communication infrastructures remains structurally dependent on access channels that are privately owned, privately governed, and unilaterally alterable. From a political-economy perspective, this produces a deep asymmetry: platforms are able to observe society at scale, while society can observe platforms only intermittently, conditionally, and often on terms set by the platforms themselves. This policy paper sets out DIGIPOL's recommendations within this context.

Introduction

Public-interest digital research expanded significantly during the 2010s, making it increasingly possible to generate public knowledge about platform governance decisions, political speech online, and tracking societal phenomena such as pandemics. Yet obtaining that knowledge has frequently been mediated through interfaces, APIs, and commercial tools that can be restricted, redesigned, priced, or withdrawn. Since 2022, the European Commission has increasingly developed legislation to moderate the online sphere. Enforcement actions have also been initiated – for example, in April 2024, the European Commission opened formal proceedings against Meta under the Digital Services Act (DSA), including in relation to the non-availability of an effective third-party, real-time monitoring tool following the deprecation of CrowdTangle. These developments underscore a core lesson of the DIGIPOL project: Observability cannot rest on discretionary and revocable interfaces if it is to support democratic oversight in a durable way.

Legally, the European Union has gone further than any other jurisdiction in recognizing this problem and beginning to institutionalize a response. The DSA, which became applicable on 17 February 2024, created an initial observability architecture through Article 40 researcher-access provisions, statements of reasons for moderation decisions, its Transparency Database, systemic-risk assessment duties, and independent audit requirements for especially consequential digital services – termed very large online platforms (VLOPs) and very large online search engines (VLOSEs). The Commission had already designated the first VLOPs and VLOSEs in April 2023, thereby triggering enhanced obligations for those services. In January 2024, the Commission sent formal requests for information to designated services specifically concerning their obligation to provide eligible researchers with access, without undue delay, to publicly accessible platform data. In February 2025, it also released a Research API for the DSA Transparency Database to facilitate programmatic analysis of moderation data. Taken together, these

measures represent an important shift: observability is no longer treated solely as a matter of voluntary corporate transparency, but as a legitimate subject of public law, institutional design, and enforcement.

Yet, as of June 2025, the EU framework is better understood as a first layer of observability than as a complete observability regime. The law created disclosure duties and access routes, but it did not yet establish a fully stable public-interest ecosystem capable of withstanding platform-side instability. DIGIPOL demonstrates that the relevant policy problem is not only whether access exists in principle, but whether it is sufficiently usable, continuous, and insulated from private disruption to support cumulative research, replication, and timely public scrutiny. On the regulatory side, the implementation difficulties visible across the Union by mid-2025 show that coherent political commitment remains necessary: in May 2025, the Commission referred five Member States — Czechia, Spain, Cyprus, Poland, and Portugal — to the Court of Justice for failing to properly empower a Digital Services Coordinator (DSC).

The broader argument of this policy paper is that platform governance must move not only from transparency as ‘disclosure’ to transparency as ‘effective observability’, but also that it must move from procedure to infrastructure. DIGIPOL’s central lesson is that observability must be treated as a public-interest capacity capable of surviving the withdrawal of APIs, the disappearance of proprietary tools, and the broader enclosure and monopolisation dynamics of the platform economy. Article 40 is a valuable starting point because it moves data access towards a regulated pathway. However, what is ultimately needed is an observability ‘regime’: one that combines enforceable access rights with stable institutions, shared standards, archival continuity, and independent research capacity.

This policy paper outlines some steps towards this direction. The necessity for this is evidenced by high-salience domains such as migration - a DIGIPOL focus. In both the United States and the European Union, policymakers have highlighted how migration narratives can exploit fear and anxiety, undermine trust, and fuel social division. Much of this high-impact speech on high-stakes topics now takes place within algorithmically shaped environments. The construction of a robust observation ecosystem of such environments is therefore not only a methodological imperative for relevant social science research, but also a fundamental democratic necessity if citizens are to retain meaningful oversight over the forces shaping their societies and politics.

1. Recognize platform observability as a public-interest infrastructure

Policy should be to treat observability not as a form of platform transparency, even if legislated, but as a public-interest infrastructure necessary for democratic oversight. The introduction of API restrictions experienced in the 2020s is not a dependency of individual research projects, but rather, it reflects a wider structural dependence upon platforms in order to regulate platforms. In the EU, the DSA has already created an initial regulatory process through access rules and also a Transparency Database for moderation decisions. However, these compose an initial entanglement of compliance devices rather than a holistic observability regime. While policy needs to protect technological development, a more mature approach would support dedicated observatories, with stable funding, with shared standards, and preservation functions that survive platform-side changes in APIs, interfaces, or terms. In short, public observability should withstand private decisions of monitoring interfaces provided by platforms, even those provided with access rules through legislation.

2. Move from discretionary access to enforceable access pathways

A robust observability regime should reduce dependence on access channels that platforms can unilaterally restrict, redesign, price, or withdraw. Article 40 of the DSA is a positive example precisely because it converts access from a matter of private permission into a regulated process. For

jurisdictions beyond the EU, the principle should be similar: establish a statutory access route, designate a competent authority, and require platforms to publish clear points of contact, data-access information, and reasons for refusal. However, an imbalance must nevertheless exist between centralized control of data and decentralized assessment.

3. Require public data access to be genuinely usable in practice

A formal entitlement to platform data is not sufficient if access is partial, historically shallow, or subject to opaque rate limits and restrictive terms. Public-interest research requires data access that is complete enough to support analysis that is replicable. The DSA is explicit that qualifying researchers should receive access to publicly accessible data “without undue delay” and, where technically possible, to real-time data. However, this should be comprehensive and proportionately historical, and sufficiently comprehensive, even where data is compartmentalized into distinct or algorithmically delineated public spheres. This should be neutral to data type, whether it be text, audio, video, or otherwise. Interfaces should be usable, and safeguards should be instated against throttling, quota design, or historical limitations that defeat the research purpose. In principle, public-data access should be treated as a substantive obligation, not a symbolic one.

4. Standardize workflows, timelines, and decision formats across jurisdictions

Even well-designed access rights can fail if the surrounding procedures are opaque, fragmented, or excessively burdensome. Moreover, access for public-interest research should not be contingent on political will of individual jurisdictions, disenfranchising researchers working on local research outside of the jurisdictions of strong regulatory bodies such as the EU, or disenfranchising researchers with weak informal ties. A robust observability regime should therefore require standard application templates, indicative processing timelines, amendment procedures, reasoned decisions, and common minimum service standards so that access does not depend on informal relationships, national variability, or institutional capacity asymmetries alone.

5. Require platforms to publish data catalogues and access metadata before requests are made

Researchers cannot make proportionate requests if they have no reliable picture of what data exists, in what structure, under what retention conditions, and through which access modalities it may be provided. One of the most practical features of the new Article 40 regime is therefore the expectation that VLOPs and VLOSEs make DSA data catalogues available. These catalogues should be updated regularly, including when emerging risks are identified through ad hoc risk assessments or audit reports. Platforms subject to heightened scrutiny should maintain public-facing catalogues describing the existence, scope, format, and retention horizon of research-relevant datasets, together with points of contact, access routes, and known constraints. This would significantly reduce uncertainty, improve necessity and proportionality assessments, and make access procedures less adversarial and less wasteful for researchers, regulators, and platforms.

6. Use privacy-preserving and security-preserving access modalities as an enabling mechanism, not as a reason for default refusal

A mature observability regime should not treat privacy, confidentiality, trade secrets, and security as automatic vetoes. It should instead require institutions to match the access modality to the risk profile

of the data. Regulators should develop tiered models of access, ranging from interfaces or APIs for lower-risk public data to secure data rooms or remote enclaves for more sensitive data. The objective should be proportionate access under safeguards; not blanket enclosure justified in the name of rights protection. Likewise, protection modalities should be weighed by policymakers in order to ensure that digital legislation as a whole is consistent with a public-interest observatory – in the case of the EU, this implies ensuring that GDPR privacy protections are in-line with DSA requests.

7. Shift the policy benchmark from transparency to auditability

Observability is weak if researchers can see data once but cannot revisit, compare, or reproduce findings over time. This is why documentation, versioning, and continuity should be treated as core policy requirements rather than technical extras. The European Commission's own work on the DSA Transparency Database points in this direction: it launched a Research API in February 2025 to enable programmatic analysis of content moderation data. Yet it is necessary to ensure that sufficient historical 'memory' is maintained for cumulative, longitudinal public-interest research on changing moderation patterns, crisis periods, or election cycles. A stronger policy approach would therefore require longer archival horizons where lawful, stable schema documentation, persistent identifiers, machine-readable codebooks, and changelogs whenever platforms alter interfaces, categories, recommender features, or moderation-related systems that affect data retrieval or analysis. Without that, the observability is reproduced through instability – over which legislation has no effect.

8. Enable explanation systems to support oversight

Statements of reasons and transparency databases should be treated as oversight instruments, not merely as compliance repositories. The DSA already requires statements of reasons to contain clear and specific information, including the grounds on which content was removed or restricted, and the Commission's Transparency Database makes those statements available almost in real time for search and download. These are important advances. However, the next policy step should be to improve the usability of explanation systems for regulators, researchers, journalists, and civil society. That means discouraging formulaic or templated justifications, stabilizing categories across time, and adding metadata that make context and legal basis easier to interpret at scale. The Commission's own release of the Research API and open-source analysis tools shows the value of building research-facing functionality into transparency systems. The same principle should guide future legislation.

9. Invest in independent observatories and research capacity, not only in case-by-case access

Access procedures are necessary but insufficient. Public-interest observability also requires institutions that can preserve expertise, maintain archives, support replication, and absorb shocks when platform interfaces or access conditions change. Independent technology research is under pressure from shrinking funding, limited access, and growing professional risk for researchers. Governments, regulators, and funders should support consortia, observatories, and common infrastructures that pool legal, technical, and methodological capacity rather than leaving individual researchers or research organizations to negotiate alone. In the EU context, bodies such as the European Centre for Algorithmic Transparency (ECAT) can help coordinate, but they should complement rather than replace an independent observatory ecosystem. The aim should be continuity: multi-year support, common methods, secure shared environments, and infrastructures that remain usable when any one platform, tool, or research partnership disappears.

10. Build regulator capacity in high-salience domains (such as migration) as stress tests for observability

An observability regime will only be as strong as the institutions charged with implementing it. The DSA already assumes a high-capacity supervisory model. Regulators need specialist staff able to assess research applications, data-security safeguards, proportionality, and platform compliance, and they need mechanisms for cooperation across regulatory silos. High-salience domains such as migration and elections should be used as practical stress tests for whether observability systems are fit for purpose. Institutions and researchers alike have stressed the escalating impact of deliberate disinformation, hate speech, and significantly harmful narratives, and have called for coordinated, multi-level responses grounded in reliable, accessible facts. Regulators need particular capacity to enable researchers to analyze speech and platforms, inform democratic debate, to anticipate panic, and to reduce societal harm fueled by algorithmically targeted online content.

DIGIPOL is an EU-funded project hosted by the European University Institute's Robert Schuman Centre for Advanced Studies and the Harvard Kennedy School of Government's Shorenstein Center on Media, Politics and Public Policy.