



MJRC

MEDIA & JOURNALISM
RESEARCH CENTER

Who Owns the Answer: Pluralism and Power in the Automated Information Order

by Marius Dragomir

May/2026

Table of contents

01 Executive Summary and Methodology

| | |
|-------------------------|----|
| Purpose of the study | 01 |
| Key concepts introduced | 01 |
| Key findings | 02 |
| Methodology | 03 |

05 Chapter 1. Why Ownership Mapping is No Longer Enough

| | |
|---|----|
| 1.1 A scene from 2026 | 05 |
| 1.2 What the media pluralism-mapping tradition achieved | 06 |
| 1.3 What the tradition increasingly misses | 07 |
| 1.4 The licensing market as an early warning | 08 |
| 1.5 The argument of this study | 09 |

11 Chapter 2. Literature Review: Three Waves of Scholarship on Information Power

| | |
|--|----|
| 2.1 Orientation | 11 |
| 2.2 The first wave: ownership, concentration and media capture | 11 |
| 2.2.1 Philosophical foundations and measurement consequences | 12 |
| 2.2.2 Media capture as bridge concept | 13 |
| 2.3 The second wave: platforms and infrastructural dependency | 13 |
| 2.3.1 From platform dependency to automated dependency | 14 |
| 2.3.2 The platform literature's unfinished handoff | 14 |
| 2.4 The third wave: generative AI, public interest media and information integrity | 15 |
| 2.5 What the literature is yet to produce | 16 |

18 Chapter 3. The Automated Information Order: A Framework for Measuring Pluralism Beyond AI

| | |
|---|----|
| 3.1 A new vocabulary for a new order | 18 |
| 3.2 The longer-term purpose of the framework | 19 |
| 3.3 Post-authorial information and infrastructural accountability | 20 |
| 3.4 Why pluralism differs from accuracy | 21 |
| 3.5 The five layers of the Automated Information Infrastructure mapping framework | 21 |
| 3.5.1 Training data and source inclusion | 22 |
| 3.5.2 Interfaces and distribution | 23 |
| 3.5.3 Alignment as infrastructural editorial power | 23 |
| 3.5.4 Compute and models as context layers | 24 |

25 Chapter 4. What to Monitor: A Dataset Matrix for the Pluralism and Information Control Framework

| | |
|--|----|
| 4.1 Purpose and implementation of the framework | 25 |
| 4.2 Audit protocol for source visibility and publisher survey module | 29 |
| 4.3 Data governance and confidentiality | 31 |

33 Chapter 5. Applying the Framework: Seven Mapping Plans

| | |
|---|----|
| 5.1 United States | 34 |
| 5.2 United Kingdom | 35 |
| 5.3 Germany | 36 |
| 5.4 France | 37 |
| 5.5 Brazil | 38 |
| 5.6 India | 39 |
| 5.7 Netherlands | 40 |
| 5.8 Local context, mapping capacity and cross-border regulation | 41 |

43 Chapter 6. What the Data Can Be Used For

| | |
|---|----|
| 6.1 Use in regulation and oversight | 43 |
| 6.2 Use in journalism and reporting on the information environment | 45 |
| 6.3 Use in civil society engagement | 46 |
| 6.4 Use in education and public understanding | 47 |
| 6.5 Limits of the framework and indicators of success | 47 |
| 6.6 Funding, public-interest support and what the data does not by itself sustain | 48 |

50 Chapter 7. Two Information Environments in 2035

| | |
|--------------------------|----|
| 7.1 The unregulated path | 50 |
| 7.2 The regulated path | 51 |
| 7.3 The choice | 52 |

53 References

57 Contact and AI Disclosure Statement

Executive Summary and Methodology

.....

Purpose of the study

Public information environments in Europe and elsewhere are being reorganised by automated systems. Generative artificial intelligence is the most visible expression of this shift, but it is not the whole of it. Browsers, operating systems, productivity tools, search engines, voice assistants and messaging platforms are increasingly delivering public knowledge as automated answers rather than as links to identifiable journalism. The instruments that media pluralism research developed over four or so decades, mostly built around the categories of owner, outlet, audience and editorial line, were not designed for this environment. They remain necessary, but on their own they describe a shrinking share of how citizens come to know what they know.

This study sets out a framework for extending pluralism mapping into the automated environment. It introduces two analytical concepts (the Automated Information Order and the Automated Information Infrastructure), specifies what a national pluralism dataset should contain in this new environment, applies the resulting model to seven jurisdictions, and identifies the regulatory, journalistic, civic and educational uses of the data the new framework would generate. Its purpose is practical: to give national media regulators, supranational bodies and public-interest researchers a workable instrument for assessing pluralism in an information environment increasingly shaped by automated systems, while preserving the conceptual and normative commitments that have anchored pluralism research since the 1970s.

Key concepts introduced

The **Automated Information Order (AIO)** names the patterned arrangement of institutions, infrastructures, markets and dependencies through which automated systems mediate, generate, rank, summarise, personalise and monetise public knowledge. Its present form is dominated by large language models, AI search and AI assistants. Its future form may involve agents, multimodal interfaces, ambient computing or technologies not yet named. The concept is deliberately broader than “AI” to keep pluralism analysis from being tied to the vocabulary of one technological moment.

The **Automated Information Infrastructure (All)** is the operational, mappable layer of the AIO. It consists of five layers: training data, models, compute, interfaces and alignment. Three of these layers (training data, interfaces and alignment) directly produce pluralism outcomes; two (compute and models) determine the conditions under which the others operate. The framework treats All mapping as the next logical extension of the ownership-and-capture tradition, not as a replacement for it.

The **Pluralism and Information Control Framework (PICF)** is the operational instrument the study introduces. It combines a classical pluralism dataset (ownership, funding, audience, editorial conditions, regulatory environment) with an Automated Information Infrastructure dataset (training data and licensing, models, compute, interfaces, alignment), specifies a tiered feasibility model for what can be observed now and what requires statutory disclosure or audit, and provides a structured worked example of how a national coordinating body would assemble the dataset.

Given the pace at which the underlying technology is changing, the Automated Information Infrastructure component will almost certainly need to be adjusted on a rolling basis as new interfaces, model architectures and dependency structures emerge. That recurring adjustment, however, is only possible if the base is in place: the framework set out here is envisaged as the minimum infrastructure on which any future iteration can be built.

Key findings

First, classical media-pluralism instruments remain necessary but are increasingly insufficient. Ownership mapping continues to capture an important part of the information environment, but a growing share of citizens' encounter with public knowledge passes through automated systems whose editorial decisions are made through training-data choices, alignment policies and interface defaults that current pluralism instruments do not measure. The Pew Research Center's 2025 study of 68,879 Google searches by 900 US adults found that AI-generated summaries appeared in 18% of searches, that users encountering them clicked a traditional search result in 8% of visits compared with 15% without, and clicked a link inside the summary in only 1% of visits (Chapekis & Lieb, 2025). A Reuters Institute's 2026 trends survey of 280 digital leaders across 51 countries found that publishers expect search referrals to fall by more than 40% over the next three years, with Chartbeat data showing year-on-year declines already underway (Newman, 2026).

Second, the AI content licensing market is creating a stratified information economy that pluralism instruments do not currently see. Large publishers have entered direct bilateral agreements with AI providers, while local, regional, ethnic, indigenous and non-English-language journalism remains largely outside the paid licensing core (Radsch & Montoya, 2026). The result is an emerging two-tier system in which inclusion in automated information systems is being negotiated through private contracts under confidential terms, with structural pluralism consequences traditional ownership mapping cannot detect.

Third, infrastructural dependencies (compute, model access, cloud commitments and intellectual-property licensing) are increasingly becoming a locus of information power, yet they are not captured by ordinary media-ownership data. The April 2026 revision of the Microsoft–OpenAI partnership illustrates the point: it kept Microsoft as OpenAI’s primary cloud partner and retained Microsoft’s non-exclusive licence to OpenAI’s model and product intellectual property through 2032, while giving OpenAI the ability to serve its products across any cloud provider (Microsoft, 2026; OpenAI, 2026). Rather than a simple ownership story, this is a dependency structure that pluralism analysis must learn to read.

Fourth, a future-facing framework is needed. Today’s generative AI products will change. Some companies will disappear, merge or rebrand; models will become multimodal and agentic; the line between search, assistant, browser, operating system and publisher may blur further. A durable framework therefore needs to map functions and dependencies rather than product names. The five-layer Automated Information Infrastructure model introduced here (training data, models, compute, interfaces, alignment) is designed to remain useful through such changes.

Fifth, the framework has been designed to be implemented now, yet with various degrees of feasibility. A tiered feasibility model distinguishes Tier 1 (observable and actionable: interfaces, source attribution, licensing inclusion or exclusion, crawler access, publisher traffic effects), Tier 2 (partially observable: ownership, strategic partnerships, compute and model dependencies, public-sector procurement) and Tier 3 (low observability: full training datasets, internal alignment processes, confidential commercial agreements). Useful regulatory action does not depend on solving Tier 3. It depends on being able to act on Tier 1 while building disclosure regimes for Tier 2 and Tier 3 over time.

Methodology

This study was conducted by the Media and Journalism Research Center (MJRC) between mid-2025 and early 2026. It builds on data and analysis MJRC produced over the same period as part of a series of consultancies commissioned by national media regulatory authorities in Europe, with those bodies’ authorisation to draw on the underlying material here. The questions those consultancies addressed (how existing pluralism instruments perform in AI-mediated information environments, what additional data credible regulatory assessment would require, and how it could realistically be collected) are also the questions this study answers. The study draws on three lines of evidence.

The first is **document analysis**. The study reviewed a structured corpus of policy documents, legal instruments and corporate disclosures relevant to the topic. The policy and legal corpus included the European Media Freedom Act (Regulation (EU) 2024/1083), the EU AI Act (Regulation (EU) 2024/1689), the Digital Services Act (Regulation (EU) 2022/2065), the Digital Markets Act (Regulation (EU) 2022/1925), national implementing measures in the seven jurisdictions covered by Chapter 5, and the operational documentation of the Media Pluralism Monitor, MAVISE database and related instruments.

The corporate-disclosure corpus included Securities and Exchange Commission filings of the major hyperscalers and AI providers, public partnership announcements, court filings in major copyright cases, and analyst and trade-press reporting where it referenced primary corporate sources.

The second is **synthesis of the available empirical literature**, with particular reliance on the Reuters Institute's 2025 and 2026 reports on generative AI and news (Newman, 2026; Simon, Nielsen & Fletcher, 2025), the Pew Research Center's 2025 analysis of AI Overview behaviour (Chapekis & Lieb, 2025), Radsch and Montoya's (2026) mapping of the AI content licensing market, a forthcoming Colom, Murray and Poblet (2026) systematic literature review (217 sources), an OSCE Representative on Freedom of the Media's 2025 report (Schiffrin et al., 2025) and the academic journal literature on platform power, recommender systems, public-service media and AI in journalism.

The third is a series of **expert interviews** with sixteen regulatory experts in Europe (n = 16), conducted under Chatham House Rules between September 2025 and February 2026. The interview panel included current and former staff of national media regulators, competition authorities, data-protection authorities and supranational bodies, selected to cover variation in regulatory mandate (media-specific, competition, data protection), in jurisdiction (EU Member States with strong, medium and weaker pluralism infrastructures, plus one non-EU comparison), and in operational seniority. The interview protocol covered four themes: the perceived adequacy of current pluralism instruments for the AI environment; concrete experience of attempting to apply existing legal tools to AI-mediated information; obstacles to the collection of Automated Information Infrastructure data within current legal mandates; and views on the priority order in which a feasible Automated Information Infrastructure dataset would be assembled. Interview material is not quoted directly in the body of the study; it informs the diagnostic claims, the design of the tiered feasibility model in Chapter 4 and the assessment of usability in Chapter 6.

The study's purpose is to offer interested parties, including national media regulators, supranational bodies and public-interest researchers, a tool to assess pluralism in the information environment as it actually stands in 2026. It does not claim full visibility of the systems it describes. Where evidence is incomplete or contested, the study marks the gap rather than papering over it. Where data is rapidly changing, the study cites the most recent reliable source and treats the resulting figure as a snapshot rather than a fixed parameter.

Chapter 1. Why Ownership Mapping is No Longer Enough

.....

1.1 A scene from 2026

Consider a citizen in Berlin, Bucharest or Bogota who wants to understand what their government decided yesterday on a contested public question. Ten years ago, that inquiry would normally have passed through an identifiable media institution: a newspaper, broadcaster, news website, news agency or public-service outlet. The account would have been written, edited, published and, at least in principle, traceable to an organisation with a known owner, editorial line, funding structure and legal responsibility. The aggregate of such encounters across a national population was the object that media pluralism research learned to measure: who owns the outlets, what audiences they reach, how they are financed, which political and economic pressures they face, and whether citizens have access to a meaningful diversity of independent sources.

In 2026, the same citizen may increasingly pose the same question to an automated assistant, an AI-enhanced search engine, a browser, a workplace tool or a voice interface embedded in a phone or operating system. The answer arrives as a fluent synthesis. It may cite sources, or it may not. It may draw on journalism, official documents, user-generated material, proprietary databases, licensed archives, web-crawled content and model inference. It may be delivered through a search interface, browser, messaging service, productivity suite or operating-system layer. It may be shaped by training data, retrieval arrangements, ranking logic, safety policies, commercial agreements and infrastructure dependencies that the citizen cannot see and that most national regulators cannot reconstruct.

The citizen is still informed. The question is through what chain of decisions. In the legacy media environment, pluralism research could not always determine whether a particular story was fair or complete, but it could map the institutions that generated and distributed it. In the emerging environment, the visible answer is separated from much of the institutional and technical apparatus that produced it. A front-end interface may be owned by one company, the underlying model by another, the cloud infrastructure by a third, the licensed source material by a network of media partners, and the alignment and ranking policies by internal teams whose decisions are not available to public scrutiny. The relevant information chain is no longer adequately described by the categories of owner, outlet, editor and audience alone.

This study argues that the change should not be understood only as an “AI problem”. Generative AI is the current and most visible expression of a broader shift: the emergence of an Automated Information Order (AIO).

The AIO is a technological and institutional order in which automated systems mediate, organise, generate, rank, summarise, personalise and monetise public knowledge. Its present form is dominated by large language models, AI search and assistants. Its future form may involve agents, synthetic media systems, multimodal interfaces, ambient computing or technologies not yet named. The point of the concept is to keep the analysis from becoming trapped in today's terminology. The regulatory question is not how to map ChatGPT, AI Overviews, Gemini, Claude or Perplexity as isolated products. It is how to map the durable infrastructure through which automated systems are likely to shape the public sphere after the current generation of tools has changed.

This study calls that mappable infrastructure the Automated Information Infrastructure (AII). The AII is the operational layer of the AIO: the training data, models, compute infrastructure, interfaces, alignment systems, licensing arrangements and contractual dependencies through which automated information is produced and delivered. It is not a substitute for the media system. It sits around, above and increasingly inside it. Journalism continues for now to supply much of the verified information on which democratic life depends. But the path from journalistic production to public knowledge is being reorganised by automated infrastructure. The analytical move proposed here is therefore continuous with the best traditions of media pluralism research. The field moved from counting newspapers to analysing ownership, from ownership to capture and platform dependency, and must now move from media institutions alone to automated information infrastructure. Ownership mapping is still needed, but by itself it describes a shrinking share of the power relations that determine how citizens come to know what they know.

1.2 What the media pluralism-mapping tradition achieved

Before describing what is missing, it is important to recap what the existing tradition has achieved. The modern media pluralism literature begins with a democratic concern that remains valid: the power to shape public opinion is a form of political power, and concentrated control over that power can undermine democratic self-government (Baker, 2006).

The ownership-pluralism tradition turned that concern into a research programme. Noam (2009, 2016) established a benchmark for systematic measurement across sectors and over time. Baker (2006) articulated the normative case for ownership dispersion as a democratic condition, not merely an economic preference. Doyle (2002) connected media ownership to convergence, competition and regulatory design, showing that measurement choices are themselves normative choices. Karppinen (2013) pushed the field beyond simple counting of outlets, arguing that pluralism must be understood as a question of structural power relations in the media environment. Pickard (2020) extended this argument by reframing journalism itself as a public good whose sustainability cannot be left to market mechanisms alone.

Institutional monitoring translated these insights into operational practice. The Centre for Media Pluralism and Media Freedom's Media Pluralism Monitor uses a methodology built around four dimensions (Fundamental Protection, Market Plurality, Political Independence and Social Inclusiveness) and approximately 200 variables per country (Media Pluralism Monitor, 2025). Reporters Without Borders mapped ownership and concentration risks across national media and the European Audiovisual Observatory's MAVISE database provides structured information on audiovisual media services and video-sharing platforms in Europe, including service providers, jurisdictions and, where available, final ownership (European Audiovisual Observatory, 2026). The Media and Journalism Research Center's Media Influence Matrix[1], State Media Monitor (Dragomir, 2025), Media Capture Monitoring Report[2] and related projects sit within the same tradition while extending its attention from formal ownership to political and economic influence (Dragomir, 2019).

This tradition has three strengths that any new framework would need to preserve. First, it provides a shared vocabulary: external and internal pluralism, horizontal and vertical concentration, beneficial ownership, editorial independence, media capture and public interest. Second, it has generated repeatable datasets, which allow longitudinal and comparative analysis. Third, it has institutional credibility: regulators, courts, journalists, civil society organisations and researchers can use it. The tradition continues to function as designed. The problem is that the environment it was designed to describe is no longer the whole information environment.

1.3 What the tradition increasingly misses

The first strain appeared with digital platforms. Search engines, social networks, video platforms and messaging systems became dominant routes through which people encountered news. Ownership of a newspaper or broadcaster remained important, but control over distribution, ranking, recommendation and monetisation increasingly sat elsewhere. Platform scholarship showed that publishers still produced much of the news, while platform companies increasingly controlled the channels through which people accessed it (Nielsen & Ganter, 2022; van Dijck, Poell & de Waal, 2018). Regulatory responses such as the Digital Services Act (Regulation (EU) 2022/2065) and Digital Markets Act (Regulation (EU) 2022/1925) reflect this recognition. They treat very large platforms and gatekeepers as infrastructural actors with obligations that go beyond ordinary market participation.

The second strain is deeper. Automated information systems do not only distribute journalistic content. They transform it: they summarise, synthesise, answer, translate, personalise and sometimes hallucinate. They may reduce a complex investigative article to a sentence, cite a source without sending traffic to it, integrate an outlet's archive into a model, or exclude an entire language community because its material is poorly represented in training data. The problem is no longer only who distributes media content, but who controls the systems that transform media and other knowledge into automated answers.

[1] See more at <https://mediainfluencematrix.org>.

[2] See more at <https://journalismresearch.org/media-capture-monitoring-report/>, and <https://ipi.media/report-assessing-media-capture-and-emfa-implementation-in-2025/>.

Three limitations of the existing tradition follow. First, the unit of analysis is changing. Classical ownership mapping identifies entities and ownership links. The AIO often works through a mix of equity, licensing, application programming interfaces, cloud commitments, model access, revenue sharing and infrastructure dependency. The April 2026 amendment to the Microsoft–OpenAI partnership illustrates this. Microsoft remained OpenAI’s primary cloud partner and retained a non-exclusive licence to OpenAI’s model and product intellectual property through 2032, while OpenAI gained the ability to serve its products across any cloud provider (Microsoft, 2026; OpenAI, 2026). This agreement reveals a contractual and infrastructural dependency, renegotiated as the market evolves.

Second, reach measurement is also changing. Circulation, viewership, audience share and web traffic remain useful indicators for media outlets. But AI interfaces may mediate information without creating a measurable visit to a source. The Pew Research Center’s 2025 analysis of 68,879 Google searches by 900 US adults found that about 18% produced an AI-generated summary; users who encountered such a summary clicked a traditional search result in 8% of visits, compared with 15% when no AI summary appeared, and clicked a link inside the summary in only 1% of visits (Chapekis & Lieb, 2025). These figures are not a universal measure of news consumption, but they demonstrate why outlet-level audience metrics are no longer enough.

Third, editorial influence is also changing. Editorial independence has traditionally referred to the relationship between owners, managers and newsrooms. In automated systems, output is shaped by training data, retrieval arrangements, ranking logic, alignment policies, safety rules, product defaults and commercial incentives. These processes are editorial in effect, even if they are not editorial in professional identity. They determine what is visible, what is excluded, what is summarised, what is treated as authoritative and what is made frictionless for users (Helberger et al., 2018).

1.4 The licensing market as an early warning

The emerging AI content licensing market is one of the clearest indicators that information power is being reorganised. Since 2023, major publishers have entered into agreements with AI companies. OpenAI announced a global partnership with Axel Springer in December 2023 that included summaries of selected content from Politico, Business Insider, BILD and WELT and use of the publisher’s content in OpenAI systems (OpenAI, 2023). It announced a multi-year global partnership with News Corp in May 2024, covering access to current and archived content from major News Corp mastheads (News Corp, 2024; OpenAI, 2024). Other deals have been reported with publishers and agencies including the Associated Press, Financial Times, Le Monde, Prisa, The Atlantic and others.

The policy significance of these deals is not simply that publishers are being paid, but that inclusion in automated information systems is being negotiated through private contracts, often under confidential terms. Radsch and Montoya (2026) document that direct bilateral deals are concentrated among large publishers with brand recognition and legal capacity, while local, regional, ethnic, indigenous and non-English-language journalism remains largely outside the paid licensing core. A forthcoming literature review similarly finds that bilateral licensing arrangements do not solve the problems of pluralism, information integrity or the vulnerability of small and medium-sized public interest media (Colom et al., 2026).

The result is a possible two-tier journalism economy. Large publishers may be licensed, visible and remunerated. Smaller publishers may be scraped without meaningful compensation, technically blocked, or excluded from AI outputs because they are commercially marginal, linguistically underrepresented or difficult to license at scale. This goes beyond market competitiveness, becoming a pluralism problem. Which media are licensed, retrievable, cited, summarised or ignored will shape the informational diet of citizens, a layer that traditional ownership mapping is unable to capture.

1.5 The argument of this study

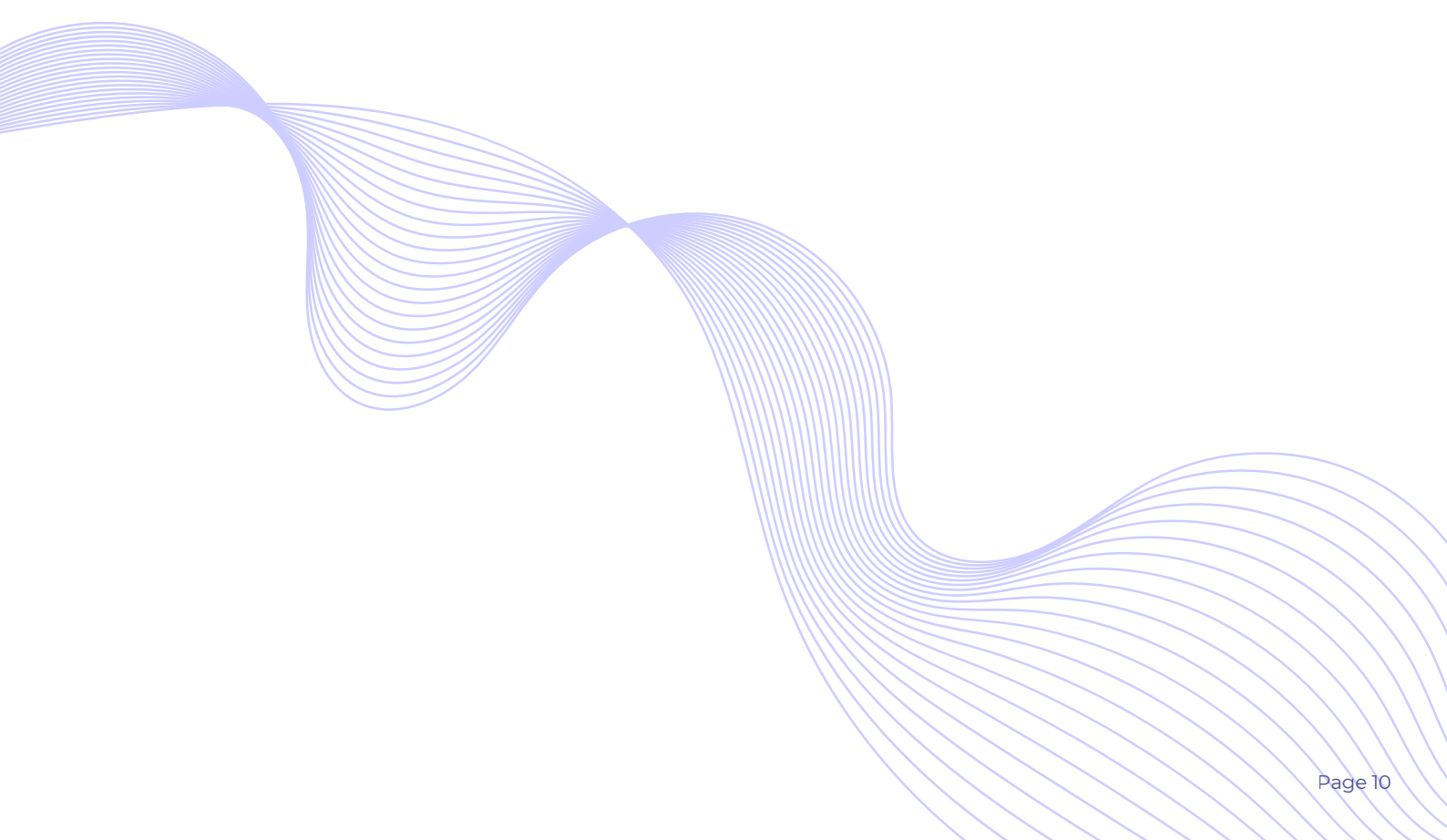
This study advances four linked claims.

First, traditional media pluralism mapping remains necessary. The outlets whose ownership it tracks still produce journalism, set agendas, investigate power and maintain professional standards. The concentrations and capture risks identified by ownership and media capture research remain democratic concerns.

Second, ownership mapping is no longer sufficient. In the AIO, informational power is exercised through infrastructure as well as ownership: through interfaces, model access, licensing markets, data flows, compute dependencies and alignment systems. Third, pluralism research must not be absorbed into a narrow AI-governance vocabulary. AI safety, copyright, competition law and technical auditing all matter, but none of them alone answers the pluralism question, which is: who has power over the conditions under which citizens encounter public knowledge?

Fourth, a new pluralism framework must be built for the longer term. Today's generative AI products will change. Some companies will disappear, merge or rebrand; models will become multimodal, agentic and embedded in devices; thus, the line between search, assistant, browser, operating system and publisher may blur further. A durable framework therefore needs to map the Automated Information Order rather than the vocabulary of one technological moment.

The remainder of the study develops this argument. Chapter 2 reviews the relevant literature across three waves: ownership pluralism, platform power and AI-mediated information. Chapter 3 defines the AIO and AIi and explains why a new framework must be future facing rather than tied to current AI terminology. Chapter 4 translates the concept into a tiered mapping model. Chapter 5 applies the model to seven jurisdictions to show how the research can be conducted in different national contexts. Chapter 6 discusses how the data the framework would generate can be used in regulation, journalism, civil society and education. Chapter 7 closes with a scenario-based stress test of what the information environment may look like in a decade with and without effective pluralism mapping.



Chapter 2. Literature Review: Three Waves of Scholarship on Information Power

.....

2.1 Orientation

The central question of this study (who shapes what citizens come to know, through what systems and under what constraints?) has developed through three waves of scholarship. The first wave focused on media ownership and structural pluralism. The second focused on platform power and algorithmic mediation. The third, still forming, addresses automated information systems and the political economy of generative AI. These waves are not replacements for one another. Each identifies a new layer of power while leaving earlier layers in place.

This chapter is diagnostic rather than exhaustive. It identifies what each wave contributed and left unresolved, and why a new framework is needed. Its conclusion is that the field already possesses many of the concepts required for the next step (ownership, capture, gatekeeping, platform dependence, infrastructural power, public interest media, epistemic welfare and information integrity) but lacks an integrated measurement framework for the AIO.

2.2 The first wave: ownership, concentration and media capture

The first wave begins with the idea that pluralism is not only a matter of content variety, but of the distribution of communicative power. The foundational works of this tradition were introduced briefly in Chapter 1: Baker (2006) on the democratic case for ownership dispersion; Doyle (2002) on convergence and cross-media control; Noam (2009, 2016) on the systematic measurement of concentration across sectors and over time; Karppinen (2013) on pluralism as a question of structural power relations. The Media Pluralism Monitor (2025) operationalises these insights through indicators across fundamental protection, market plurality, political independence and social inclusiveness.

The concept of media capture extended the first wave by showing that formal ownership is only one pathway of influence. Capture operates through state advertising, regulatory pressure, media licensing, politicised public media governance, ownership transfers, litigation, selective subsidies and informal political-economic relationships (Dragomir, 2019). This move is particularly important for the present study because it shows how pluralism research can move from entities to relationships. The Automated Information Infrastructure requires a similar move: from who owns an outlet to who controls the conditions under which information is transformed and delivered.

2.2.1 Philosophical foundations and measurement consequences

The first wave matters not only because it produced data, but because it established why pluralism is a democratic rather than merely economic concept. A competition authority may ask whether prices rise, entry is blocked or market power is abused. A pluralism framework will ask whether citizens encounter a sufficiently independent and diverse set of sources for democratic deliberation. The two questions overlap, but they are not the same. A media market can be competitive in a narrow economic sense and still fail pluralism if the available voices are politically dependent, socially narrow or structurally aligned with the same interests.

This distinction becomes even more important in the AIO. Automated systems may be offered at zero monetary price to users. They may increase convenience, reduce search costs and improve access to information. A purely consumer-welfare analysis may therefore see little harm. But pluralism analysis asks whether the resulting informational pathway narrows the diversity of sources, reduces the visibility of public interest journalism, undermines editorial accountability or makes citizens dependent on infrastructures whose values and incentives are not publicly contestable. Pickard (2020) makes this argument explicitly for journalism: treating news as an ordinary market product underestimates the democratic infrastructure it sustains.

The measurement consequence is also important. Classical ownership research never measured pluralism perfectly. It used proxies: ownership shares, audience shares, revenue concentration, cross-media holdings, regulatory safeguards and editorial-independence indicators. These proxies were imperfect but useful because they were stable, comparable and linked to plausible theories of power. The Automated Information Infrastructure mapping should proceed in the same spirit. It should not wait for perfect access to model weights, full training datasets or internal alignment documents. It should identify stable proxies for infrastructural control: interface reach, licensing inclusion, source attribution, model dependency, cloud dependency and output behaviour across repeated tests.

2.2.2 Media capture as bridge concept

Media capture is the bridge between classical ownership analysis and Automated Information Infrastructure analysis. Capture research showed that the owner of an outlet is not always the only actor who matters. Governments can capture media through advertising, taxation, public procurement, licensing, litigation, surveillance and selective regulatory enforcement. Business actors can capture media through debt, distribution control, informal political alliances or ownership intermediaries. In such environments, formal ownership data may be accurate and still insufficient (Dragomir, 2019).

The same logic applies to Automated Information Infrastructure. An AI company may not own a publisher, but it may control whether the publisher's content is crawled, licensed, cited, summarised, monetised or excluded. A cloud provider may not own an AI lab, but it may shape the lab's cost structure, scaling capacity and product distribution. A platform may not edit a story, but it may integrate an AI answer box that captures the user's attention before the story is reached. Capture research therefore provides a methodological lesson: pluralism mapping must follow influence where influence actually operates, not only where older legal categories expect it to operate.

2.3 The second wave: platforms and infrastructural dependency

The second wave emerged because the platform era exposed a gap in classical pluralism tools. Media organisations still produced journalism, but Google, Facebook, YouTube, Twitter/X, TikTok, Apple News and messaging platforms increasingly controlled access, visibility and monetisation. Nielsen and Ganter (2022) describe platform power as relational: publishers are empowered by platforms and dependent on them at the same time. Van Dijck, Poell and de Waal (2018) analyse platforms as institutions that organise public values and social practices, not merely as technical intermediaries.

This scholarship reframed gatekeeping. Traditional gatekeeping theory asked how editors selected news. Platform-era gatekeeping asked how ranking, recommendation, moderation and data-driven advertising systems shaped visibility (Helberger et al., 2018). It also raised the question of dependency. Publishers could gain reach through platforms, but on terms set by platforms. The history of referral traffic, algorithm changes, demonetisation, platform grants and news-product closures shows how media organisations became vulnerable to changes in private infrastructures they did not control. Schiffrin et al. (2025), in the OSCE Representative on Freedom of the Media's 2025 report, document how this vulnerability has been carried over and intensified in the AI environment.

EU regulation reflects the second wave. The Digital Services Act (Regulation (EU) 2022/2065) imposes obligations on online platforms and very large online platforms, including systemic risk assessment, transparency and access to data for vetted researchers. The Digital Markets Act (Regulation (EU) 2022/1925) identifies gatekeepers that provide core platform services such as search engines, operating systems, web browsers, virtual assistants, cloud services and online advertising services. These instruments are not media pluralism laws in the narrow sense, but they recognise that digital infrastructures can exercise public-sphere power.

Platform scholarship and regulation, however, mostly treat platforms as distributors or intermediaries of content. The AIO goes further: automated systems increasingly become answer systems. They do not merely decide which source a user sees first. They may decide whether the user sees a source at all.

2.3.1 From platform dependency to automated dependency

The platform era also produced a cautionary policy lesson. Publishers accepted the platform bargain because the immediate benefits (reach, traffic, social discovery, search referrals, advertising opportunities) were visible. The costs (audience dependency, advertising-market capture, algorithmic volatility, weakened direct relationships with readers, migration of data value to intermediaries) were slower to appear. By the time many publishers recognised the structural imbalance, the infrastructure had become difficult to exit. The AIO may repeat this pattern in a compressed timeframe. Licensing deals, AI-search integrations, summarisation products and publisher tools may offer short-term revenue or technical advantages. But if the market develops through bilateral deals and proprietary interfaces, publishers may again become dependent on infrastructures whose terms they cannot set. This is why the framework proposed here treats early mapping as preventive regulation: the aim is to identify dependency structures while they are still forming.

2.3.2 The platform literature's unfinished handoff

Platform scholarship changed the object of media regulation by showing that distribution infrastructures can exercise editorial and economic power. Its unfinished handoff lies in the move from distribution to generation. Platform analysis often assumed that a user clicked from a platform to a source. Automated answer systems weaken that assumption. The user may receive the answer without clicking, may not know which source was used, or may trust the synthesis more than the underlying reporting (Brantner et al., 2025). The five-language probing study by Brantner and colleagues found systematic differences in how AI-powered search engines surface political news depending on the language of the query, a finding that illustrates the new pattern.

This change does not make platform concepts irrelevant. Gatekeeping, dependency, self-preferencing, data extraction and infrastructural power remain central. But they must be re-specified. The gate is no longer only a ranking position in search results or a recommendation feed. It may be the decision to cite, not cite, retrieve, summarise, license, block, paraphrase or refuse. That is the conceptual step the present study takes.

2.4 The third wave: generative AI, public interest media and information integrity

The third wave is still unstable. It includes AI ethics, copyright, platform governance, competition law, journalism studies, information integrity and public interest media research. Its empirical base is developing rapidly, and some claims are necessarily provisional. Nevertheless, several clear themes are emerging. The most systematic synthesis of the available literature to date, upcoming in 2026 (Colom et al., 2026), draws on 217 sources from systematic and snowball searches; Mansell et al. (2025), in a Nordicom volume on information ecosystems and troubled democracy, complements that synthesis with a state-of-knowledge review covering news media, AI and data governance; Schneiders and Stark (2025) develop the case for what they term “epistemic governance” of platformised news ecosystems.

The first emerging theme is extraction. Large-scale scraping of journalistic content, infrastructure strain from AI crawlers and asymmetric bargaining power between GenAI providers and media organisations have emerged as recurring themes (Radsch & Montoya, 2026; Colom et al., 2026). The AI licensing market risks reproducing the structural dependency of the platform era through new tollbooths: licensing intermediaries, bilateral deals and voluntary compensation arrangements controlled by the largest technology actors (Radsch, 2024; Radsch & Montoya, 2026).

The second is traffic displacement. The relationship between AI summaries and publisher traffic remains contested, and figures vary by market, query type and methodology. What is clear is that AI-generated summaries and answer engines weaken the old bargain in which search engines indexed content and returned users to websites. Pew’s click-through findings, cited in Chapter 1, show lower clicking when AI summaries are present (Chapekis & Lieb, 2025). Reuters Institute reporting in 2026 found that surveyed media leaders expected search traffic to decline by more than 40% over the next three years (Reuters Institute for the Study of Journalism, 2026). These findings should not be treated as final causal estimates, but as early evidence of a structural shift in referral logic.

The third is licensing stratification. The content licensing market is forming before public rules have stabilised. Bilateral deals may bring revenue to major publishers, but they are not scalable for the whole information ecosystem. They may also encourage a patronage dynamic in which AI companies decide which publishers are worthy of payment and integration (Radsch & Montoya, 2026). Collective licensing, statutory remuneration, levy proposals and data commons have therefore become part of the policy debate (Keller, 2025; Molitorisz & Attard, 2024).

The fourth is opacity. Training datasets, retrieval sources, model behaviour, alignment policies and ranking rules are not transparent enough for pluralism assessment. The EU AI Act (Regulation (EU) 2024/1689, Article 53) creates obligations for providers of general-purpose AI models, including technical documentation and summaries of training content, but these obligations are not designed as a complete pluralism-monitoring system. The DSA may be more useful for some information-environment questions because it treats large platforms as systemic intermediaries, but it too was not designed specifically for automated news synthesis (Seipp et al., 2024).

The fifth is information integrity. Literature on public interest media emphasises that journalism is not simply content. It is an institutional practice that verifies facts, contextualises events, holds power to account and maintains epistemic standards (Pickard, 2020). Van den Bulck et al.'s (2025) idea of public service media as a contributor to “epistemic welfare” is especially relevant: public media help create conditions in which citizens can acquire justified beliefs and knowledge. Automated systems may support this function, but they may also bypass the institutions that sustain it. Sjøvaag (2024) traces the implications for the business of news in the AI economy, while Abiri (2024) analyses generative AI itself as a form of digital media subject to the regulatory and normative frameworks developed for earlier media forms. Pattison, Ricks and Wihbey (2025) argue that AI-driven search may reshape democracy, economics and human agency in ways that are not yet fully visible to existing measurement instruments.

2.5 What the literature is yet to produce

The literature now recognises many of the symptoms: extraction, disintermediation, platform dependency, licensing asymmetry, opacity and traffic displacement. What it has not yet produced is a pluralism framework that is both conceptually broad and operationally realistic.

The conceptual gap is that the third wave is often framed too narrowly as a debate about AI. AI systems are part of a broader order of automated mediation. The next interface may not be called an AI assistant; it may be a browser, an operating system, a workplace agent, a school tutor, a civic-service bot, a synthetic news channel or an augmented-reality layer. A pluralism framework should therefore map functions, dependencies and control points rather than product names.

The operational gap is that regulators cannot map everything. Full training datasets are often proprietary, massive, constantly changing and legally contested. Alignment systems are dynamic and partly internal. Cross-border cloud infrastructure lies outside many national regulators' direct reach. A realistic framework should therefore distinguish between what can be observed now, what can be partially observed through disclosure and investigation, and what remains opaque but can be tested through audits, sampling and risk assessment (Carlini, 2025).

To conclude, a pluralism-oriented third wave should avoid at least four errors. First, it should not treat AI as a self-contained technical field. The issue is not simply model accuracy or safety. It is the distribution of power over public knowledge. Second, it should not treat copyright licensing as the whole solution. Copyright can support compensation and bargaining, but pluralism also concerns visibility, source diversity, accountability and democratic inclusion. Third, it should not assume that transparency automatically produces accountability. Dataset summaries, model cards and voluntary reports can be useful, but only if they are tied to enforceable obligations, audit rights and institutional capacity. Finally, it should not design a framework that only wealthy regulators can use. Local and national authorities need a minimum viable dataset that can be collected with realistic resources. The framework should be ambitious in concept but tiered in practice.

Chapter 3. The Automated Information Order: A Framework for Measuring Pluralism Beyond AI

.....







3.1 A new vocabulary for a new order

The term Automated Information Order is introduced here to name a durable shift in the organisation of public knowledge. An order is more than a technology. It is a patterned arrangement of institutions, infrastructures, markets, norms and dependencies. The broadcast order was organised around spectrum, licences, public-service obligations, advertising markets and national audiences. The platform order was organised around search, social networks, recommendation systems, programmatic advertising and data extraction. The Automated Information Order should be organised around automated systems that generate, summarise, rank, translate, personalise and transact information on behalf of users.

The term is deliberately broader than “AI”. Artificial intelligence is the current enabling technology, but public policy should not have to reinvent its pluralism framework every time the dominant technical label changes. The same pluralism questions will arise if today’s large language models are replaced by agentic systems, synthetic knowledge engines, personal data vaults, neuro-symbolic systems or other successors. The relevant question is whether automated systems control the pathways through which citizens encounter public knowledge.

The Automated Information Infrastructure is the measurable object within this order. It consists of five layers: training data, models, compute, interfaces and alignment. These layers do not operate separately. Training data affects model outputs; compute determines who can build and deploy systems at scale; interfaces determine reach; alignment shapes acceptable answers; licensing and contractual arrangements connect the layers. Pluralism risks emerge from their interaction.

Overview of core concepts

| Concept | Definition | Why it matters for pluralism |
|---|--|---|
|  Automated Information Order (AIO) | The broader technological and institutional order in which automated systems mediate, generate and organise public knowledge. | Keeps the study future-facing and prevents the framework from being tied only to current AI products. |
|  Automated Information Infrastructure (All) | The measurable infrastructure of the AIO: data, models, compute, interfaces, alignment, licensing and contractual dependencies. | Provides the object that policymakers, regulators and researchers can map. |
|  Pluralism and Information Control Framework (PICF) | The mapping framework proposed in this study, integrating classical media pluralism datasets with All datasets. | Translates concepts into a feasible regulatory tool. |
|  Post-authorial information | Information whose immediate user-facing form is shaped by automated systems rather than encountered as the work of an identifiable author or outlet. | Identifies the user experience of automated information and the accountability gap it creates. |
|  Tiered feasibility | A model that distinguishes between All layers that can be observed now and those that require statutory disclosure or audits. | Makes the framework usable today, even where full transparency is not yet legally available. |
|  Public-interest media inclusion | An assessment of whether public interest journalism, including local, regional, minority-language and investigative outlets, is included in automated information systems. | Directly tests whether automated systems narrow or broaden the diversity of voices reaching citizens. |

3.2 The longer-term purpose of the framework

The most important conceptual change in this study is the explicit long-term framing. This framework is not designed only for the first decade of generative AI, but for the period after the current AI debate has moved on. Regulatory systems often lag behind technical change. Media ownership rules developed for print and broadcasting had to be adapted to cable, satellite, digital platforms and streaming. The same pattern will repeat if pluralism regulation treats current AI assistants as the final object.

The framework should therefore be read as a future-proof pluralism instrument. In the short term, it may help regulators map generative AI systems, licensing markets and AI search. In the medium term, it may be used to map agentic systems that select, purchase, summarise or negotiate information on behalf of users. In the long term, it may be useful to map whatever automated mediation layer succeeds today's AI products. Its categories are functional: source inclusion, infrastructure dependency, interface dominance, output governance and visibility. Those functions are likely to persist even if the underlying technologies change.

This longer-term framing also matters politically. The current AI debate is crowded with arguments about model safety, copyright, labour displacement, industrial policy, national security and innovation. While all those are important, the pluralism question risks being lost inside that debate. The Automated Information Order is a concept used by this study to make visible the architecture of public knowledge under automated mediation.

3.3 Post-authorial information and infrastructural accountability

The phrase “post-authorial” can be misunderstood. It does not mean that authors will disappear: journalists, experts, officials, witnesses, researchers and citizens continue, and will probably continue, to produce the knowledge that automated systems recombine. Rather, the term describes the user’s experience of information in AI-mediated environments. As search shifts from lists of sources toward “answer engines,” the immediate answer is often encountered less as the work of a named author or institution than as a system response. Authority shifts from the byline to the interface (Gillespie, 2014; Narayanan Venkit et al., 2025; Ofcom, 2025).

This has two consequences for accountability. On the one hand, the user may not know which institution should be trusted, challenged or corrected. If an automated answer misrepresents a policy decision, is the relevant accountable actor the publisher whose story was summarised, the AI provider, the search engine, the model developer, the retrieval system or the user-facing interface? On the other hand, the institutions that created the original information may lose the reputational and economic link to the public. A news organisation can perform costly verification while an automated system captures the user’s attention and trust.

The framework therefore treats accountability as infrastructural. An accountable AIO would preserve traceability between automated outputs and underlying sources; make source diversity visible; allow errors to be corrected; disclose major dependencies; and give regulators enough data to assess systemic patterns. Accountability does not require every answer to become a legal document. It instead requires the system to maintain enough provenance and auditability that public knowledge is not detached from responsibility.

3.4 Why pluralism differs from accuracy






A common response to concerns about automated information is to focus on factual accuracy. Yet, while accuracy is essential, but pluralism is broader. A system can be accurate in the narrow sense and still reduce pluralism if it consistently cites the same dominant sources, ignores local journalism, excludes minority languages, or presents contested public questions as settled administrative facts. A pluralistic system is not one that treats all claims as equally valid, but one that gives citizens access to a range of independent, relevant and accountable sources, especially where public questions are contested.

The distinction matters for regulation. Technical benchmarks may test whether a model answers factual questions correctly, but pluralism audits need to test source diversity, attribution, contestation, local relevance and public interest visibility. A model that answers “what did parliament pass yesterday?” with a correct summary but cites only a government press release is less pluralistic than one that also surfaces independent reporting, opposition reactions and expert analysis. The answer is not wrong, but the informational pathway is narrow. The Automated Information Infrastructure framework should therefore be connected to, but not absorbed by, AI safety testing. Safety evaluation asks whether a system causes harm through false, dangerous or unlawful outputs. Pluralism evaluation asks whether the system contributes to the conditions of democratic knowledge.

3.5 The five layers of the Automated Information Infrastructure mapping framework

These new realities that have emerged in today’s communication space show why single-layer analysis is inadequate. A model trained on diverse sources may still produce narrow outputs if its interface privileges a few partners. A publisher may be licensed but rarely cited. A national regulator may know which media companies operate domestically but not which AI interface citizens use to access national news. On the other hand, a regulator may collect AI-interface data but miss the licensing market that determines which publishers appear inside it. The Automated Information Infrastructure mapping framework part of the Pluralism and Information Control Framework (PICF) proposed in this study attempts to address these complexities. It consists of five layers, three of them directly relevant for information pluralism (training data and licensing, interfaces and distribution and alignment and editorial policy)

The Automated Information Infrastructure mapping framework: the datasets needed for the new wave of pluralism regulation and research

| Layer | Function | Pluralism question | Main evidence sources |
|---|--|--|---|
| Layers directly relevant to the new pluralism regime | | | |
|  Training data and licensing | Determines what material enters, is excluded from or is made retrievable by automated systems. | Which media, languages and communities are included, licensed, blocked, scraped or absent? | Licensing disclosures, publisher surveys, crawler data, litigation, AI Act summaries. |
|  Interfaces and distribution | Deliver automated information to citizens through search, assistants, browsers, messaging and devices. | Which interfaces dominate civic information, and how do they attribute or hide sources? | Usage surveys, DSA data access, audits, app-store and platform data. |
|  Alignment and editorial policy | Shapes what systems will say, refuse, emphasise or treat as authoritative. | Whose values, sources and incentives shape the answer, and is the process accountable? | Documentation, structured probing, incident reports, regulator findings. |
| Context layers | | | |
|  Models | Process and generate outputs based on trained parameters and retrieval inputs. | Which models mediate public information, who controls them, and how dependent are users on closed systems? | Model catalogues, vendor disclosures, procurement data, technical documentation. |
|  Compute | Provides the cloud, chip and energy infrastructure on which models and interfaces run. | How concentrated is compute provision, and what dependencies does it create? | SEC filings, competition cases, infrastructure disclosures, sovereign cloud strategies. |

3.5.1 Training data and source inclusion

Training data is the layer through which source inclusion begins. It matters for pluralism because it determines what kinds of knowledge are represented in the statistical and retrieval foundations of automated systems. For journalism, the key questions are whether public interest media are included, whether inclusion is licensed or extracted, whether small and local outlets are represented, whether non-English and minority-language content is present, and whether data from high-quality journalism is distinguishable from low-quality or synthetic content (Abiri, 2024).

This is also the least straightforward layer to regulate. Full training datasets are often proprietary, too large to inspect directly and partly composed of material whose legal status is contested. The EU AI Act’s requirement for providers of general-purpose AI models to make available summaries of training content (Regulation (EU) 2024/1689, Article 53) is a useful entry point, but it is not a full pluralism audit. A regulator cannot infer source diversity, language balance or public-interest representation from a high-level summary alone. The mapping framework therefore treats training data as a low-observability but high-importance layer.

The practical approach is indirect. Regulators and researchers can map licensing agreements, crawler access, blocked domains, publisher participation in marketplaces, litigation, and output citation patterns. These indicators do not reveal the full training corpus, but they reveal the political economy of inclusion.

3.5.2 Interfaces and distribution

Interfaces are the most immediately mappable layer. Citizens encounter automated information through products: search engines, AI assistants, browsers, voice assistants, messaging systems, productivity tools, learning platforms, navigation tools and operating systems. Some of these products are already within the scope of existing EU digital regulation. The DMA definition of core platform services includes online search engines, operating systems, web browsers, virtual assistants, cloud computing services and online advertising services (Regulation (EU) 2022/1925, Article 2). The DSA applies to online platforms and very large online platforms, including obligations relevant to systemic risk and data access (Regulation (EU) 2022/2065).

For pluralism, the key questions are: which interfaces are used for news and civic information; what source attribution they provide; whether they link to original media; whether they privilege licensed partners; whether they suppress or demote certain categories of content; and whether users understand when they are receiving an automated synthesis rather than an editorial product.

Interfaces should therefore be the first tier of regulatory mapping. Unlike training data and alignment, they are externally observable. Regulators can survey usage, run structured queries, require transparency reports, and examine traffic effects through collaboration with publishers and researchers.

3.5.3 Alignment as infrastructural editorial power

Alignment is the most conceptually difficult layer to map. In AI governance, alignment is often discussed in relation to safety, helpfulness, refusal behaviour or compliance with developer policies. In pluralism analysis, alignment should be understood more broadly as a form of infrastructural editorial power. It shapes what an automated system will say, refuse, emphasise, tone down, cite or treat as authoritative.

This does not mean alignment teams are equivalent to editors in a newsroom. They do not commission reporting or follow journalistic codes. But their decisions can have editorial effects. A model may give different levels of confidence to official sources, commercial sources, public interest media, partisan outlets or user-generated content. It may avoid certain political questions, over-rely on English-language material, flatten local context or convert contested topics into falsely neutral summaries (Brantner et al., 2025).

Because alignment systems are dynamic and internal, regulators cannot map them in the same way they map ownership. The feasible approach combines disclosure, documentation, incident reporting and independent testing. Structured probing can identify output patterns across languages, topics and time. This is not a perfect method, but neither is audience measurement or content analysis in legacy media. The issue is not perfect visibility, but whether pluralism-relevant patterns can be observed, repeated and assessed.

3.5.4 Compute and models as context layers

Compute and models are context layers because they shape the conditions under which the core pluralism layers operate. Compute concentration matters because only a small number of firms can provide the cloud, chips, data-centre capacity and energy infrastructure required for frontier-scale systems. Model concentration matters because a small set of models may underlie many products, even when user-facing interfaces differ.

These layers should not distract regulators from the more directly pluralism-relevant questions of source inclusion and interface visibility. But they cannot be ignored. A national media regulator may not be able to access directly relevant information about cloud infrastructure, but it needs to know whether the country's public information environment depends on a handful of foreign providers or companies aligned with various groups of interests. A competition authority may not assess editorial diversity, but it can investigate whether cloud, model and distribution dependencies create lock-in.

Chapter 4. What to Monitor: A Dataset Matrix for the Pluralism and Information Control Framework

.....






4.1 Purpose and implementation of the framework

The purpose of the framework is not to map every technical detail of every AI system. That would be impossible and unnecessary. Its purpose is to produce a dataset that would allow regulators, researchers and public interest actors to answer five questions: which automated interfaces mediate news, civic and political information in a given jurisdiction; which media sources are licensed, cited, retrievable, blocked or absent from those systems; which firms control the relevant models, cloud infrastructure and distribution channels; which legal and contractual arrangements shape the inclusion of journalism in automated systems; and which pluralism risks are visible only when classical media data and Automated Information Infrastructure data are analysed together.



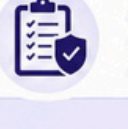


The framework is therefore a leverage-based mapping model. It does not promise total system visibility, but rather identifies what can be observed now, what requires statutory disclosure, and what must be assessed through audits and sampling.

That being said, as part of a renewed regulatory architecture, the categories in the classical pluralism datasets should not be weakened. Automated Information Infrastructure mapping should sit alongside them, not replace them. Without classical data, regulators cannot know whether automated systems amplify already concentrated media markets or compensate for them by exposing users to more diverse sources.

Pluralism and Information Control Framework component one: classical pluralism datasets




| Category | What the dataset should contain | Why it remains necessary |
|--|--|--|
|  Ownership structures | Beneficial ownership, corporate control, cross-ownership, cross-border holdings, annual ownership changes. | Journalism production remains institutionally located; ownership still shapes editorial independence and market power. |
|  Funding and revenues | Advertising, subscriptions, public funding, state advertising, philanthropy, platform payments, AI licensing revenues (where disclosed). | Sustainability and dependency cannot be assessed without revenue structure. |
|  Audience and reach | Television, radio, print, digital reach, subscriptions, traffic, trust and consumption surveys. | Outlet reach remains necessary for assessing public influence and market concentration. |
|  Editorial independence and professional conditions | Editorial statutes, newsroom governance, labour conditions, safety, legal threats, unionisation, self-regulation. | Structural independence cannot be inferred from ownership alone. |
|  Regulatory environment | Concentration rules, merger tests, transparency obligations, public-service governance, press-freedom safeguards. | Regulation determines what pluralism risks can be addressed. |

Pluralism and Information Control Framework component two: Automated Information Infrastructure (AII) datasets

| All category | What to collect | Principal contributors | Main limits |
|--|--|---|---|
| Layers directly relevant to the new pluralism regime | | | |
|  Training data and licensing | Publisher-AI deals; marketplace participation; crawler-blocking data; opt-out and robots.txt practices; litigation; AI Act training-content summaries. | Publishers, press associations, AI providers, copyright offices, competition authorities, AI Office, researchers. | Confidential contracts; incomplete training disclosures; cross-border data. |
|  Interfaces and distribution | AI assistant/search/browser usage; civic-information use; attribution and linking patterns; click-through rates; default settings in operating systems and browsers. | Media regulators, national statistical offices, DSA data access, publishers, independent auditors. | Limited platform data access; changing interfaces; sampling bias. |
|  Alignment and output governance | Published model specifications; content policies; refusal and ranking patterns; structured tests on civic queries; incident reports. | AI providers, vetted researchers, regulators, civil society, academic labs. | Internal opacity; version changes; limited external observability. |
| Context layers | | | |
|  Models | Major models accessible in the jurisdiction; open/closed status; deployment in public services and newsrooms; procurement dependencies. | AI Office, procurement bodies, competition authorities, public-service media, researchers. | Vendor opacity; model rebranding; API layers. |
|  Compute | Cloud providers, data-centre capacity, public-sector AI cloud procurement, energy implications, cross-cloud dependency. | Competition authorities, telecom/energy regulators, procurement authorities, cloud providers. | Extraterritoriality; commercial confidentiality; lack of sectoral data. |







For the Automated Information Infrastructure part of the framework to be implemented, a tiered feasibility model is needed as a main operational correction as not all layers can be mapped with equal precision. In the short term, regulation should begin with Tier 1 because it yields actionable evidence quickly. Tier 2 should be coordinated with competition and procurement bodies. Tier 3 requires disclosure and audit regimes and should be treated as an area for progressive institutional development rather than immediate full visibility.

A tiered feasibility model for the implementation of the Pluralism and Information Control Framework

| Tier | What belongs here | Feasibility | Regulatory value | Recommended tools |
|--|---|-------------|------------------|--|
|  <p>Tier 1: observable and actionable</p> | Interfaces, source attribution, licensing inclusion/exclusion, crawler access, publisher traffic effects. | High | High | Surveys, audits, DSA data access, publisher reporting, EMFA pluralism assessments. |
|  <p>Tier 2: partially observable</p> | Ownership, strategic partnerships, compute and model dependencies, public-sector procurement. | Medium | Medium to high | Competition inquiries, procurement transparency, merger review, market studies. |
|  <p>Tier 3: opaque but critical</p> | Full training data, alignment systems, internal ranking logic, proprietary evaluation methods. | Low | High | AI Act documentation, audit rights, structured probing, incident reporting, research access. |









The European Media Freedom Act (EMFA) is particularly important because it explicitly recognises media pluralism and editorial independence as distinct from ordinary competition analysis. Article 22 requires Member States to provide rules for assessing media market concentrations that could significantly affect media pluralism and editorial independence, including effects on the formation of public opinion and taking account of the online environment (Regulation (EU) 2024/1083, Article 22). The Automated Information Infrastructure framework can help operationalise that online-environment clause for automated mediation.

Use of EU legal instruments in the Pluralism and Information Control Framework

| Instrument | Relevant hook | How it supports Automated Information Infrastructure pluralism mapping | Limitation |
|---|--|---|---|
|  Digital Services Act | Systemic risk assessment, transparency reporting, data access for vetted researchers. | Can support research on interface visibility, recommender systems, civic-information risks and source distribution. | Not designed specifically for AI answer systems or media pluralism. |
|  Digital Markets Act | Gatekeeper rules for core platform services including search, browsers, operating systems, virtual assistants and cloud. | Helps identify distribution bottlenecks and prevent self-preferencing or lock-in in automated interfaces. | Focuses on contestability and fairness, not pluralism as such. |
|  European Media Freedom Act | Media-market concentration assessments and consideration of online environment and public-opinion formation. | Provides the strongest media-pluralism hook for assessing automated mediation and platform/media dependencies. | National implementation and institutional capacity will determine practical effect. |
|  AI Act | General-purpose AI documentation, training-content summaries, systemic-risk obligations. | Provides entry points for dataset and model documentation. | Transparency summaries are unlikely to be granular enough for pluralism assessment. |
|  Competition law | Market investigations, merger review, abuse of dominance, cloud/AI partnership scrutiny. | Can expose dependency structures and contractual concentration. | Competition harm and pluralism harm are related but not identical. |
|  Copyright and related rights | Licensing, text-and-data-mining reservations, remuneration models. | Can support compensation and consent mechanisms. | Does not address visibility, diversity or source integrity by itself. |

A national media regulator need not perform all tasks itself. Its role is coordination, standard setting and public synthesis. The model should be a distributed data collection under a standard methodology, producing a comparable risk picture.

Overview of data collection processes for the Pluralism and Information Control Framework

| Contributing entity | Classical pluralism contribution | Automated Information Infrastructure contribution |
|--|--|--|
|  National media regulator | Ownership, licensing, media concentration, public-service governance. | Coordinates AII pluralism dataset; runs interface/source audits; integrates EMFA assessment. |
|  Competition authority | Merger data, market studies, cross-ownership analysis. | Maps AI/cloud/platform partnerships, gatekeeper dependencies and licensing market concentration. |
|  Data protection authority | Data-use and privacy compliance affecting media and users. | Reviews data-processing issues in AI interfaces and public-sector AI deployments. |
|  National statistical office | News consumption, trust, media-use surveys. | Adds AI-use questions, including news/civic/political information categories. |
|  Press and publishers' associations | Sector economic data, revenue and distribution information. | Provides confidential licensing, crawler, traffic and referral data. |
|  Public service media | Audience data, governance, funding and editorial-independence information. | Reports AI tool adoption, source partnerships and automated distribution arrangements. |
|  Universities and research institutes | Independent analysis, methodology and content studies. | Conducts structured probing, source-diversity audits and longitudinal studies. |
|  Civil society and journalist associations | Press freedom, safety, labour and editorial independence. | Collects incidents, source visibility concerns and impacts on small/local media. |

4.2 Audit protocol for source visibility and publisher survey module

For the Pluralism and Information Control Framework to work, a source-visibility audit should be simple enough to repeat and rigorous enough to compare. The coordinating body should define a standard query set covering national politics, local government, public health, elections, courts, public spending, climate events, cultural policy and minority communities. Queries should be run in all major national languages and, where relevant, regional or minority languages. The same query should be tested across major interfaces at regular intervals.

For each output, auditors should record whether the system provides an answer; whether it cites sources; whether citations are clickable; whether sources are media, government, academic, commercial or user-generated; whether local or regional sources appear; whether public service media appear; whether the answer encourages the user to consult original sources; and whether the output contains factual errors or misleading omissions.

The audit should not attempt to certify truth in every case, but rather to identify patterns. If a system consistently cites government sources but rarely independent media on contested policy questions, that is a pluralism risk signal. If it cites English-language sources for national-language queries, that is a language-representation risk signal. If it cites licensed partners disproportionately, that is a market-structure risk signal. If it gives long answers without links, that is an attribution risk signal.








A publisher survey should be short, confidential and repeated annually. It should ask whether the outlet has: signed direct AI licensing agreements; participated in a marketplace or intermediary; received requests from AI companies; blocked or allowed major AI crawlers; detected crawler traffic; received traffic from AI interfaces; changed its robots.txt or technical access policies; experienced attribution failures; used AI tools in the newsroom; and adopted internal AI guidelines.

The survey should classify outlets by size, geography, language, ownership type and public-interest function. This is essential because aggregate publisher data can hide inequality. A licensing market that covers most national audience share may still exclude most outlets. A public-interest pluralism analysis should therefore ask how many local, regional, community, minority-language and investigative outlets are included, not only how many users are reached by licensed content.

Survey results should be published in aggregated form. Individual contract values need not be disclosed, but the regulator should be able to report categories: e.g., no contact from AI companies; contacted but no deal; direct licence; marketplace participant; blocked crawlers; unable to monitor; uncertain. Such categories would already produce information that does not currently exist in most jurisdictions.

A regulator-facing framework should convert concepts into indicators. The following indicators are intended as a starting point for an annual or biennial Automated Information Infrastructure pluralism module. They are deliberately modest: each can be collected through a combination of public data, confidential surveys, structured audits or statutory disclosure. The indicators should be interpreted together. A country may have diverse media ownership but low automated source diversity. Another may have high AI adoption but strong public service media visibility. Another may have few domestic AI actors but effective disclosure and audit capacity. **Collected and analysed together, these indicators can provide a risk diagnosis useful for information control and pluralism analysis and regulation.**

Overview of indicator design by layer

| Layer | Primary indicator | Secondary indicators | Suggested risk question |
|--|--|--|---|
|  Interfaces | Share of population using major automated interfaces for news/civic information. | Frequency of use; age/language breakdown; default integration in devices and browsers. | Are a few interfaces becoming unavoidable gateways to public information? |
|  Source attribution | Share of audited civic answers that cite identifiable media sources. | Clickable links; source diversity; local/regional/minority-language representation. | Are citizens shown where information comes from, and are diverse sources visible? |
|  Licensing | Share of major national, local and minority-language publishers covered by AI licensing or marketplace arrangements. | Deal type; exclusivity; remuneration model; opt-out/blocking status. | Does the licensing market produce inclusion or stratification? |
|  Traffic effects | Change in search/AI referral traffic for publishers by size and type. | Click-through from AI summaries; direct visits; subscription conversion. | Are automated interfaces weakening the economic pathway from discovery to journalism? |
|  Alignment/output governance | Existence of public model specifications and output policies relevant to civic information. | Version logs; incident reports; external audit results. | Are output rules transparent enough for democratic scrutiny? |
|  Model dependency (context layer) | Main models used by dominant interfaces and public-sector systems. | Open/closed status; provider jurisdiction; API dependence. | Is public information mediated by a narrow set of foreign-controlled models? |
|  Compute dependency (context layer) | Cloud providers used for major AI systems and public deployments. | Data-centre location; energy commitments; exit options. | Is the public information layer dependent on concentrated infrastructure? |

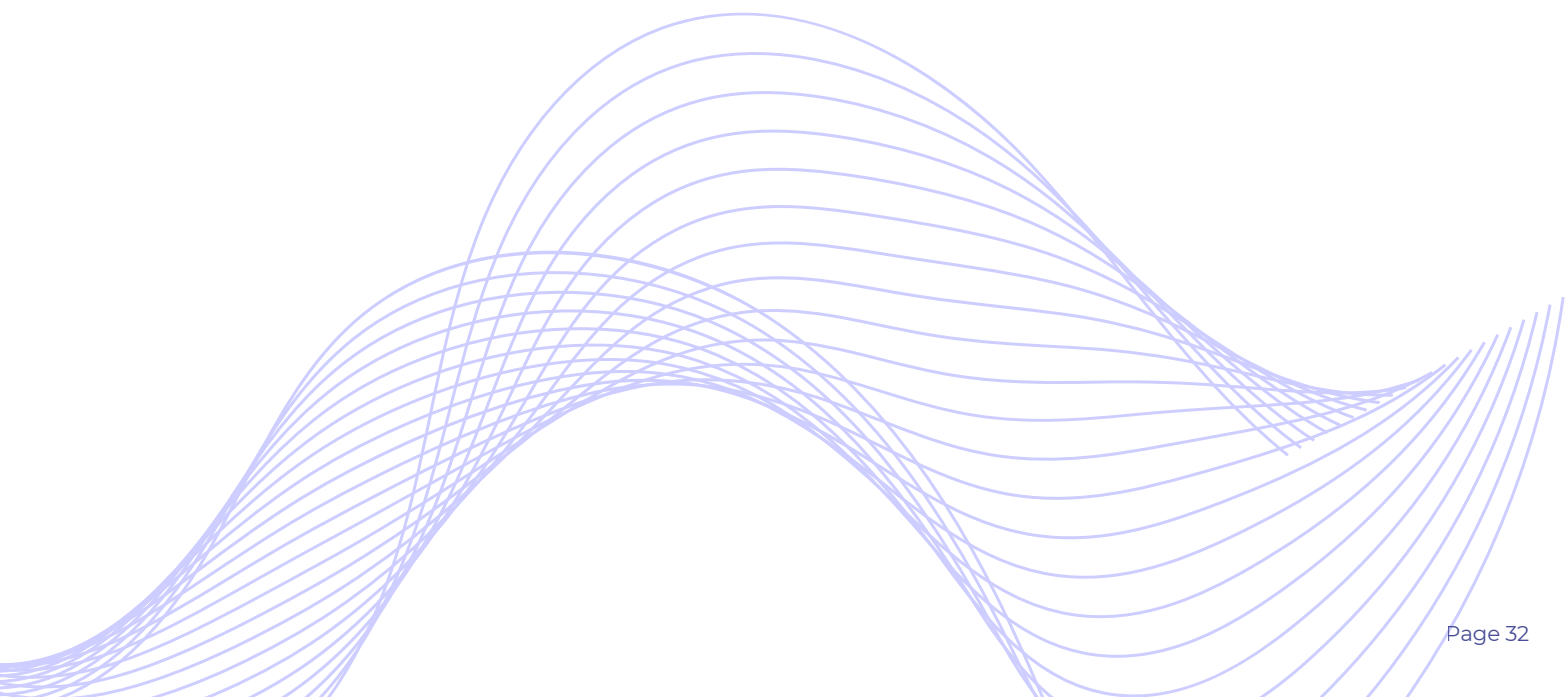
4.3 Data governance and confidentiality

Automated Information Infrastructure mapping will require data that some actors regard as commercially sensitive: licensing terms, crawler logs, traffic impacts, revenue shares, cloud contracts and system prompts. A credible framework therefore needs a data-governance model. Not all data collected by regulators may be published in raw form. Some can be collected confidentially and reported in aggregated categories. For example, a regulator may publish the share of local publishers with licensing arrangements without disclosing individual contract values. It may report the proportion of audited AI answers citing local sources without publishing every test query if doing so would enable gaming.

This is especially important for small publishers. If Automated Information Infrastructure mapping depends only on data voluntarily supplied by large platforms and large publishers, it will reproduce the asymmetries it is meant to diagnose. Public funding may be needed to help smaller outlets provide crawler data, traffic analytics and licensing information.

One major limitation to the effective implementation of the framework is that many of the required disclosures do not yet exist. That is not, however, a reason to delay. Media pluralism instruments often began with imperfect data and improved as monitoring became institutionalised. The first generation of Automated Information Infrastructure mapping can rely on surveys, public announcements, output audits, publisher coalitions and existing platform data. As legal obligations develop, the dataset can become more granular.

Waiting for perfect law would repeat the platform-era mistake: regulators would act only after markets, defaults and dependencies had hardened. Early mapping has value even when incomplete because it identifies missing data as a policy problem. A regulator that cannot determine which AI interfaces mediate civic information, or which local media are visible in automated answers, has learned something important: the information environment is already outside the reach of its current instruments.



Chapter 5. Applying the Framework: Seven Mapping Plans






.....

While Chapter 4 specified what a pluralism-and-Automated Information Infrastructure dataset should contain, this chapter applies the model to seven jurisdictions: the United States, the United Kingdom, Germany, France, Brazil, India and the Netherlands.

An important clarification on what this chapter does and does not do. The framework has been applied to these seven jurisdictions in the sense that, for each, the study has identified which national entities would be the principal contributors to each dataset block, what the consolidated dataset would reveal in that national context, and how feasible the collection appears given existing legal instruments, regulator capacity and data-availability conditions in 2026. The study has not collected the data. What the seven country sections present are mapping plans rather than completed national datasets. They are a demonstration of how the research could be conducted, organised and resourced in different national contexts. Their purpose is to make the framework concrete enough that a national media regulator, supranational body or research consortium could read the relevant section and recognise what work would actually need to be commissioned, by whom, on what timeline.





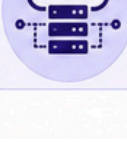
The selected cases fall into two broad categories. The United States is a source jurisdiction: many of the relevant Automated Information Infrastructure actors are headquartered or legally anchored there. The other cases are dependency jurisdictions to varying degrees: they host important media systems and regulators, but much of the Automated Information Infrastructure used by their citizens is controlled abroad. Germany, France and the Netherlands are also EU Member States and therefore relevant to DSA, DMA, EMFA and AI Act implementation. Brazil and India illustrate large democratic markets outside the EU where language, scale and regulatory capacity create distinct mapping challenges.

5.1 United States

| Dataset block | Main contributing entities | What the dataset would reveal | Feasibility |
|---|--|---|-------------------|
|  Classical pluralism | FCC, FTC, DOJ, SEC filings, academic monitors, unions. | Ownership, advertising concentration, tech-media acquisitions, journalism labour and local news decline. | Medium |
|  Interfaces | Platform disclosures, Pew/academic panels, browser/search market data. | Use of AI search and assistants for civic information; role of US-headquartered platforms as global distributors. | Medium |
|  Licensing | Publisher associations, court filings, public deal announcements, AI licensing intermediaries. | Which major publishers are paid/integrated; whether local and ethnic media are included. | Medium |
|  Alignment/output | Academic audits and civil society testing. | How US-focused civic queries are answered and sourced across major systems. | Medium-low |
|  Models and compute | FTC/DOJ inquiries, cloud-market data, public company filings. | Dependency structures among AI labs, hyperscalers and enterprise distribution. | Medium |




The United States is the most important source jurisdiction because many global AI, cloud, search, browser and platform firms are headquartered there. A mapping plan executed in the US would therefore reveal not only domestic pluralism risks but the upstream structures affecting other countries. The difficulty is that US media pluralism regulation is more fragmented and less explicitly pluralism-oriented than the EU framework. Competition, consumer protection and copyright litigation may, in practice, supply more usable data than media-specific regulation.

5.2 United Kingdom

| Dataset block | Main contributing entities | What the dataset would reveal | Feasibility |
|---|---|--|--------------------|
|  Classical pluralism | Ofcom, CMA, ICO, BBC, industry bodies, unions. | Media ownership, public-service media conditions, news consumption and trust. | High |
|  Interfaces | Ofcom media literacy and technology-use research; CMA market studies. | AI assistant and search use among adults and children; platform dependency. | Medium-high |
|  Licensing | Publishers, NMA, public deal tracking, AI intermediaries. | Whether UK national publishers are integrated differently from local publishers. | Medium |
|  Alignment/output | Ofcom/CMA-linked research and university audits. | Source visibility and accuracy in UK political and local-information queries. | Medium |
|  Models and compute | CMA, public procurement, cloud-market inquiries. | Relationship between AI deployment and cloud concentration. | Medium |





The UK is well positioned for Automated Information Infrastructure mapping because Ofcom and the Competition and Markets Authority already have experience with platform, media and digital-market analysis. The main risk is that AI-mediated distribution becomes a competition and innovation issue without a sufficiently explicit pluralism lens. A mapping plan executed in the UK would help connect media sustainability, public service media and AI distribution within a single analytical frame.

5.3 Germany

| Dataset block | Main contributing entities | What the dataset would reveal | Feasibility |
|---|---|--|--------------------|
|  Classical pluralism | Landesmedienanstalten, KEK, Bundeskartellamt, ARD/ZDF, research institutes. | Cross-media concentration, public-service media, regional pluralism and publisher economics. | High |
|  Interfaces | National surveys, public-service media analytics, DSA data access. | Which automated interfaces mediate German-language civic information. | Medium-high |
|  Licensing | Publisher associations, individual media groups, public deal tracking. | Whether large national publishers are licensed while regional/local media are absent. | Medium |
|  Alignment/output | German-language structured probing by universities and regulators. | Source diversity, local context and political neutrality in German-language answers. | Medium |
|  Models and compute | Bundeskartellamt, public procurement, EU AI Office data. | Dependence on US cloud/model infrastructure and European alternatives. | Medium |






Germany illustrates the value of combining strong classical pluralism infrastructure with Automated Information Infrastructure data. Existing concentration and public service media frameworks provide a baseline. The new questions a mapping plan would address are whether automated interfaces increase or reduce exposure to regional and local journalism, and whether German-language public-interest sources are visible in AI outputs.

5.4 France

| Dataset block | Main contributing entities | What the dataset would reveal | Feasibility |
|--|--|---|--------------------|
|  Classical pluralism | ARCOM, Autorite de la concurrence, CNIL, public service media, press bodies. | Ownership concentration, public-service media governance, press aid and editorial safeguards. | High |
|  Interfaces | ARCOM/CNIL research, DSA data access, usage surveys. | AI search/assistant use and attribution in French-language civic queries. | Medium-high |
|  Licensing | Publishers, AFP, public deal announcements, collective management bodies. | Whether licensing reinforces large national publishers and agencies over smaller outlets. | Medium |
|  Alignment/output | CNIL/academic audits and civic-query tests. | Treatment of French political, cultural and public-service information. | Medium |
|  Models and compute | Competition authority, public procurement, European AI initiatives. | Role of European model providers and cloud sovereignty initiatives. | Medium |





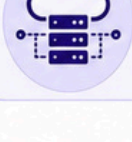
France is a relevant case study because media pluralism, cultural policy, copyright and industrial sovereignty debates intersect there. A mapping plan executed in the French context could prevent these debates from becoming siloed. A licensing deal may be good for a publisher's balance sheet and still problematic if it narrows the range of visible sources in automated outputs.

5.5 Brazil

| Dataset block | Main contributing entities | What the dataset would reveal | Feasibility |
|---|---|---|-------------------|
|  Classical pluralism | CADE, communications authorities, academic monitors, journalist associations. | Broadcast concentration, regional inequalities, government advertising and platform dependency. | Medium |
|  Interfaces | National surveys, platform data, civil-society monitoring. | Use of AI assistants and messaging platforms for news and political information. | Medium-low |
|  Licensing | Publisher associations, public deals, crawler data. | Whether Portuguese-language and regional journalism are included in licensing markets. | Low-medium |
|  Alignment/output | Civil society and academic audits in Portuguese. | How automated systems handle elections, public health and misinformation-prone topics. | Medium-low |
|  Models and compute | Competition authority, cloud procurement and data-centre data. | Dependence on foreign models and cloud infrastructure. | Low-medium |






Brazil shows why Automated Information Infrastructure mapping should include messaging and informal distribution channels as well as search and assistants. A pluralism dataset that ignored WhatsApp-mediated news or Portuguese-language model performance would miss important parts of the national information environment. The mapping plan therefore differs in emphasis from those for European cases.

5.6 India

| Dataset block | Main contributing entities | What the dataset would reveal | Feasibility |
|---|---|---|-------------------|
|  Classical pluralism | Ministry of Information and Broadcasting, TRAI, CCI, press bodies, academic monitors. | Linguistic-market segmentation, broadcast and digital concentration, legal pressures and ownership opacity. | Medium-low |
|  Interfaces | Surveys by language and region; platform data; telecom and app usage. | AI access through messaging, search, smartphones and language tools. | Medium-low |
|  Licensing | Major publishers, news agencies, court filings, AI providers. | Whether national and English-language publishers dominate automated inclusion. | Low-medium |
|  Alignment/output | Multilingual audits across major Indian languages. | Whether automated systems handle linguistic diversity and politically sensitive queries fairly. | Low-medium |
|  Models and compute | Public AI initiatives, cloud procurement, competition studies. | Dependence on foreign frontier models versus domestic language-model initiatives. | Medium-low |

India is the strongest reminder that the Automated Information Order is not a single global environment. It is multilingual, uneven and mediated through local infrastructures. A framework designed only for English-language AI search in wealthy markets would be insufficient. A mapping plan executed in India would have to be calibrated for linguistic and regional variation from the outset.

5.7 Netherlands

| Dataset block | Main contributing entities | What the dataset would reveal | Feasibility |
|--|---|---|--------------------|
|  Classical pluralism | Commissariaat voor de Media, ACM, AP, RDI, CBS, universities. | High-quality ownership, audience and public-service data in a concentrated small-language market. | High |
|  Interfaces | CBS surveys, ACM/AP research, DSA data access. | AI use in Dutch-language information seeking and public services. | Medium-high |
|  Licensing | Dutch publishers, news agencies, trade associations. | Whether Dutch-language media are included or overshadowed by English/global sources. | Medium |
|  Alignment/output | Universities and research centres. | Source diversity and language-specific reliability in Dutch civic queries. | Medium-high |
|  Models and compute | ACM, procurement bodies, EU AI Office. | Dependency on foreign models and cloud systems in a small-language democracy. | Medium |

The Netherlands is a promising pilot jurisdiction for Automated Information Infrastructure pluralism mapping because it combines strong research capacity, relevant regulators and a media system in which language representation is a visible pluralism issue. A Dutch mapping plan could test methods later adapted elsewhere in the EU.

5.8 Local context, mapping capacity and cross-border regulation

A central test of the framework is whether it generates information that policymakers and regulators can actually use. The answer is yes, but only if the framework is prioritised. Mapping full training-data composition or internal alignment rules is not realistic for most national regulators. Mapping interface use, source attribution, publisher inclusion, licensing categories and public-sector AI procurement is realistic.

In local contexts, the most useful questions are often simple. Which AI systems are people using for news and public information? Do those systems cite local media? Do they cite public service media? Do they cite national-language sources or default to English or global sources? Which local publishers have licensed their content? Which have blocked crawlers? Which have no idea how their content is used? Do public institutions deploy AI systems that rely on foreign models without transparency about sources?

These questions can be answered with modest methods: population surveys, structured query audits, confidential publisher questionnaires, procurement reviews and cooperation with universities. The resulting data would be directly useful to regulators because it would identify where intervention is needed and possible. If the problem is interface dominance, DSA and DMA tools are relevant. If the problem is exclusion of local media from licensing, competition and media-support tools may be relevant. If the problem is poor source attribution, transparency and user-interface rules may be relevant. If the problem is public-sector procurement, government purchasing rules can change.

The framework should therefore be presented not as a maximalist demand for omniscience, but as a staged capacity-building programme. In practical terms, the first year should establish the baseline, the second should identify trends, and the third should connect indicators to enforcement, funding and policy design.

Yet the most honest test of the framework's limits sits at the national border. The Automated Information Infrastructure is, in almost every dimension that matters, transnational. The compute on which it runs is concentrated in a small number of cloud providers operating across continents. The models that mediate national information environments are trained on corpora that ignore national jurisdiction and are deployed through interfaces whose alignment decisions are taken in one country and felt in many others. The licensing markets that determine which journalism reaches automated answers in Lisbon, Bucharest or Riga are negotiated, very often, in San Francisco or New York. A regulator working only within national authority can describe how this infrastructure lands in its own jurisdiction. It cannot, on its own, describe the infrastructure itself.

The implication is straightforward, even if its institutional consequences are not. Within the AIO, meaningful pluralism measurement may not be possible, and will rarely be comprehensive, without cross-border mapping arrangements that complement the national framework set out in this study. Such arrangements need not be elaborate to be useful. A coordinated assessment across EU Member States, anchored in the European Board for Media Services, would already make visible patterns that no national regulator working alone can see: which models dominate civic information across language markets; which publishers are licensed in some jurisdictions and excluded in others; which compute and alignment dependencies are shared across the entire European information environment. None of this would displace the national framework. The national level remains the level at which most regulatory authority sits, at which most public interest journalism operates, and at which citizens encounter the public sphere. But a national framework that does not connect upward to cross-border mapping will, sooner rather than later, find itself describing the local consequences of structures it cannot reach.

Chapter 6. What the Data Can Be Used For

.....

This chapter sets out how the data the framework would produce can be put to work. Rather than telling each set of actors what they ought to do, it identifies four uses to which the data lends itself: in regulation and oversight; in journalism and reporting on the information environment; in civil society engagement; and in education and public understanding. The framing is deliberate as the framework's value depends not on prescription but on whether the data it produces becomes a usable input across the institutions and practices that already shape public knowledge. The four uses overlap in practice, and the actors involved in each commonly draw on the work of the others.

6.1 Use in regulation and oversight

The most direct use of the data collected through this framework may be in the work of national media regulators, competition authorities, data-protection authorities and supranational bodies that already have responsibility for some part of the information environment. The framework does not require new legal instruments to begin generating value. Several existing instruments already provide hooks into which the data fits.

Within the European Union, the European Media Freedom Act (Regulation (EU) 2024/1083, Article 22) provides a media-pluralism basis for assessing market concentrations that affect the formation of public opinion, including in the online environment. Automated Information Infrastructure data on interface dominance, licensing inclusion and source attribution may give Article 22 assessments empirical content they would not otherwise have. The Digital Services Act (Regulation (EU) 2022/2065) provides a data-access channel for vetted researchers to examine recommender and answer systems on Very Large Online Platforms and Very Large Online Search Engines; the framework specifies the kinds of pluralism questions that data access can be used to answer. The Digital Markets Act (Regulation (EU) 2022/1925) allows competition authorities to address gatekeeping in core platform services, including search, browsers, operating systems, virtual assistants and cloud; Automated Information Infrastructure data on dependency structures and licensing-market concentration may be needed to inform that work.

The EU AI Act (Regulation (EU) 2024/1689, Article 53) provides documentation and training-content-summary obligations for general-purpose AI providers; pluralism-relevant categories (language coverage, public-interest source representation, licensing status) can be folded into the templates the AI Office develops.

Outside the EU, the picture is uneven. National media regulators may already have the legal authority to require some of the disclosures the framework specifies; some may need new instruments. Competition authorities have tools to investigate AI, cloud and platform partnerships, and several have begun to do so. The framework's contribution may be to connect these regulatory tools to the pluralism question that is too easily lost when oversight is fragmented across competition, consumer-protection, data-protection and media bodies.

Three further regulatory uses may be worth identifying. First, the data can be integrated into existing annual or biennial pluralism assessments. National media regulators that already produce pluralism reports can include Automated Information Infrastructure indicators alongside their classical data, identifying where automated mediation is amplifying or eroding the patterns the classical instruments track. Second, the data can be used in merger and concentration review where AI, platform or licensing partnerships are at stake. A merger involving a media company and an AI provider, a platform and a publisher, or a cloud provider and a model developer can be assessed for its effects on source inclusion and public-opinion formation, not only on competitive markets. Third, the data can support a public Automated Information Infrastructure pluralism register: a public list, updated annually, of major automated information interfaces used in the jurisdiction, their ownership and model dependencies, known publisher licensing arrangements, source attribution practices, and available transparency documents. Such a register is not a licensing system. It is a transparency instrument that gives citizens, researchers and regulators a baseline picture of what is currently mediating their information environment.

At EU level, the framework may be used to support a practical institutional agenda. The European Board for Media Services can develop guidance on how EMFA media-market concentration assessments should treat automated distribution, AI search, licensing markets and platform-media dependency. The Commission can clarify how DSA systemic risk and data-access provisions can support pluralism research on automated answer systems. The AI Office can design training-content-summary templates that are useful for pluralism analysis (including language categories, source types and licensing status) while respecting trade secrets and security constraints.

6.2 Use in journalism and reporting on the information environment

The framework's data may also be a substantive resource for journalism itself, in two distinct ways.

On the one hand, the data may be a subject of reporting. Automated Information Infrastructure pluralism is one of the major underreported policy stories of the current decade. Which media are licensed by which AI companies; how local and regional journalism fares in automated answers; how AI search and AI assistants reshape referral traffic and direct relationships with audiences; how public institutions are deploying AI tools without transparency about sources; how cloud and model dependencies are concentrating in a handful of foreign providers: each of these is a story, and each is currently reported in a fragmented way across competition, technology, media-business and policy beats. A consolidated Automated Information Infrastructure dataset would give journalists a coherent factual base from which to report these stories at the level of the national information environment, rather than only at the level of individual deals or products. Trade outlets such as Press Gazette and Digiday have built useful tracking practices in this area; consolidated Automated Information Infrastructure data would make the work of those who follow them substantially easier.

On the other hand, the data that the framework can generate may be used as an input into the strategic and editorial choices that media organisations make about their own engagement with the AIO. A national publishers' association considering a sector-wide approach to AI licensing benefits from knowing how the licensing market is structured, where the dependency risks lie, and what the experience of comparable jurisdictions has been. A public service broadcaster developing an internal policy on AI tools in the newsroom would benefit from knowing which models are dominant in the national market, what their alignment characteristics appear to be on civic queries, and how their sourcing patterns treat the broadcaster's own output. A small or local publisher considering whether to allow AI crawlers, block them, or join a marketplace would benefit from knowing what the inclusion-versus-exclusion patterns look like across the sector and what the traffic and visibility consequences have been for comparable outlets. Survey-based publisher data of the kind the framework specifies (aggregated and confidential where necessary) may have practical use for editorial and business decisions, not only for regulatory monitoring.

A specific subset of journalism work is worth flagging. Newsrooms increasingly need to document how their own content is being used by automated systems: crawler logs, attribution failures, summarisation errors, traffic losses, unauthorised scraping and licensing contacts. Tools at the infrastructure layer (such as the Cloudflare pay-per-crawl initiative; Cloudflare, 2025) make this technically more feasible than it was even two years ago, but technical tools alone do not address the asymmetry of bargaining power between individual publishers and global AI providers. Aggregated newsroom documentation, fed into the framework's dataset, may become one of the strongest evidentiary bases on which sectoral and regulatory responses can be built.

6.3 Use in civil society engagement

Civil society organisations (press-freedom groups, digital-rights organisations, civil liberties bodies, public interest research centres) have been among the most active observers of the AIO question. Their engagement with the framework's data may take three forms.

The first is monitoring and advocacy on specific cases. Coalitions of press freedom and digital-rights organisations are well placed to track and surface specific incidents in which the AIO produces pluralism harms: the disappearance of a local outlet from automated answers; the systematic underrepresentation of a minority-language community; the use of AI tools by public institutions without source transparency; the suppression or distortion of contested civic information. The framework's indicator-based approach makes such incidents easier to identify, document and place in a broader pattern.

The second is independent data collection where regulators do not yet have authority. Civil society can run audits of AI civic-information outputs, conduct surveys of small and minority-language publishers, and assemble incident logs that feed into the framework's dataset. Where regulators face statutory or political constraints on what they can themselves collect, civil society organisations can produce parallel evidence that is then available to regulators, journalists and researchers. The Authors Guild's tracking of the *Bartz v. Anthropic* copyright settlement is one example of civil society creating accountability data; equivalent work can be done on the pluralism dimension.

The third is safeguarding the framework itself. Any pluralism mapping infrastructure can be misused. A government could invoke source diversity to pressure independent media, define "trusted sources" in partisan ways or demand access to data that compromises journalistic independence. This risk is not hypothetical; media capture research has documented how regulatory tools can be turned against pluralism when institutions are politically dependent (Dragomir, 2019; Schiffrin et al., 2025). The framework would therefore require guarantees: institutional independence of the mapping body; public methodology; proportionate data collection; appealability of decisions affecting media visibility; avoidance of state-defined lists of approved media. Civil society is the natural watchdog of these safeguards. The framework's success hence would depend on it being used as an instrument for pluralism rather than for control, and that distinction will not be maintained without civil-society scrutiny.

6.4 Use in education and public understanding

Media literacy as an educational practice has been built, over decades, around the idea that citizens benefit from conceptual tools for understanding the media they consume. In its earlier formulations, that meant understanding ownership, recognising bias, distinguishing news from opinion, and assessing source credibility. In the AIO, the analytical task changes shape. Citizens encountering an automated answer to a political question need a different set of tools: an understanding of training data and what it means for what the system “knows”; an understanding of alignment as a process that shapes what the system will and will not say; an understanding of interface defaults and what they mean for which sources are cited; and an understanding of the corporate-web architecture within which the system operates.

The Automated Information Infrastructure framework’s structural vocabulary, with its five layers and the relationships among them, may be teachable. It can be communicated at different levels of detail to different audiences. Public service broadcasters, public libraries, school curricula, university courses and adult-education programmes may all use a version of it. The shift from media literacy in its earlier form to what might be called automated information-infrastructure literacy is the natural pedagogical extension of the framework, and it gives the data the framework would generate a wider constituency than it would otherwise have.

Researchers also have a role here, distinct from civil society and journalism. Replicable audit protocols for automated civic information (testing source attribution, factual accuracy, viewpoint diversity, language performance and stability over time) will become increasingly useful both as an evidentiary base for the framework and as a template that other research teams can adopt. Comparative cross-country work, of the kind the seven-jurisdiction structure of Chapter 5 anticipates, may become valuable in allowing the framework to mature beyond a national exercise.

6.5 Limits of the framework and indicators of success

The framework has limits. It cannot solve the business model crisis of journalism by itself. It cannot make opaque models fully transparent. It cannot guarantee that AI systems will always provide accurate or pluralistic answers. It cannot substitute for strong public service media, independent journalism, press freedom and sustainable funding.

Its purpose is narrower but essential: to make the new structure visible enough for democratic governance. Without such visibility, regulators will continue to assess media pluralism as if the public sphere were still organised primarily around outlets and owners. That would leave the most dynamic layer of information power largely unmeasured.

The success of the framework should not be measured by whether regulators can see everything because they will not. It should be measured by whether they can answer questions that they currently cannot answer. Within three years of implementation, using indicators put forward by such new mapping frameworks, a national regulator should be able to say which automated interfaces citizens use for public information; whether those interfaces cite national and local journalism; which categories of publishers are included in licensing markets; whether small and minority-language media are excluded; how public-sector AI systems source information; and where infrastructure dependencies are concentrated.

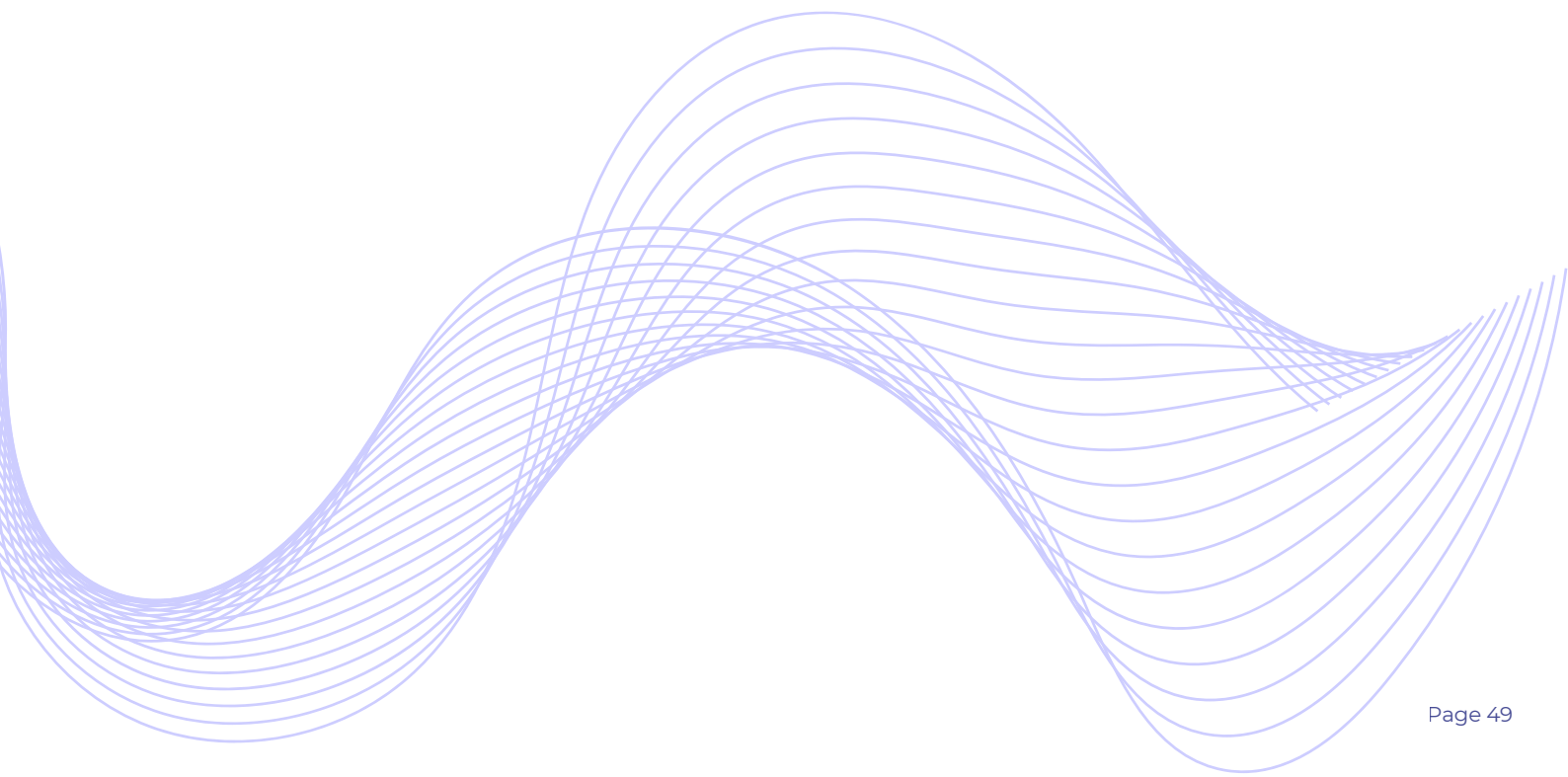
A successful framework would also change incentives. AI providers would know that source attribution and licensing inclusion are monitored. Publishers would know that their exclusion is visible as a pluralism issue, not merely a private commercial failure. Public bodies would know that procurement choices affect the public information environment. Researchers would have a common method to map the new order while citizens would gain a clearer account of how automated answers they consume are produced.

6.6 Funding, public-interest support and what the data does not by itself sustain

Mapping does not by itself sustain journalism. If Automated Information Infrastructure mapping may show, for example, that local and public interest media are structurally excluded from automated visibility and licensing markets, the data would identify a problem but would not solve it. Support mechanisms are still required: public interest media funds, innovation grants for crawler and attribution tools, collective licensing infrastructure, legal support for small publishers, and public procurement rules that favour transparent and pluralistic AI systems. Levy proposals fall into this category. The strongest version of the levy argument, as Keller (2025) puts it, is not that every act of training can be priced like a normal licensing transaction; it is that commercial automated information services derive value from a broad information commons and should contribute to the production of that commons. The design risks (governance, distribution, capture, bureaucracy, cross-border enforcement) are substantial. But the underlying problem the levy addresses is real: if value is captured at the automated aggregation layer while costs remain at the journalism-production layer, pluralism will suffer (Pickard, 2020; Sjøvaag, 2024).

Public funding should also be tied to capacity. Small media organisations often lack the technical ability to detect crawlers, analyse logs, negotiate licences or audit AI outputs. Automated Information Infrastructure pluralism policy should therefore support shared infrastructure: sector-wide analytics, model-output monitoring, legal templates, data collaboratives and technical assistance. Without such support, only large publishers will be able to participate in the automated information economy on meaningful terms.

Governments and public institutions are themselves users of automated systems, and their use is itself a pluralism question. They may deploy AI assistants for public services, education, health communication, legal information or citizen engagement. These deployments can affect pluralism if they rely on opaque retrieval systems, poor source attribution or narrow supplier ecosystems. Public sector procurement should therefore be part of any Automated Information Infrastructure mapping. The relevant questions are: which models and cloud providers are used in public information services; whether source materials are documented; whether public service media and official sources are appropriately distinguished; whether citizens are told when they are interacting with automated systems; whether the system can explain why it cites a source; and whether there is an appeal or correction mechanism for errors. Public sector deployments can also set standards: if governments require transparent source attribution, audit logs and public interest data safeguards in their own procurement, they create market incentives for better systems. On the other hand, if governments adopt opaque automated tools without pluralism safeguards, they normalise the very infrastructure problems this study identifies.



Chapter 7. Two Information Environments in 2035

.....

7.1 The unregulated path

In 2035, on the unregulated path, a citizen asks a personal civic agent what their government decided overnight. The answer is concise, personalised and confident. It summarises the decision, explains likely effects on the user's household, and offers to draft a message to a representative. It cites three sources: one government page, one large national publisher, one general knowledge database. It does not mention the local investigative outlet that first uncovered the issue, because that outlet is not part of the agent's preferred retrieval network. It does not link to a minority-language community newspaper that has covered the policy for months, because that source is not licensed and rarely appears in the model's retrieval index. It does not show the disagreement among experts, because the interface is optimised for resolution rather than deliberation.

Nothing illegal has necessarily happened. No editor has been ordered to suppress a story. No newspaper has been bought by an oligarch. No broadcaster has lost its licence. Yet the citizen's informational environment has narrowed, and this narrowing has occurred through infrastructure: data inclusion, retrieval ranking, licensing markets, interface defaults and alignment design. It is the pluralism risk of an unregulated AIO. Not censorship or ownership concentration in their classical forms, but the unseen organisation of public knowledge through systems that make some sources more retrievable, some answers more convenient and some forms of journalism less visible.

Traditional media may still exist on this path. Public service broadcasters may still broadcast. Newspapers may still publish. Independent outlets may still investigate. But the public's route to their work is increasingly mediated by automated agents whose decisions are not publicly contestable. Audience share no longer describes influence. A local outlet may be read by few directly but used by many agents; or it may produce vital reporting that agents never retrieve. A major publisher may receive licensing revenue but lose its direct relationship with audiences. A public service outlet may be cited often but stripped of context. The licensing market has become an epistemic gate, with structural effects on which voices reach the public, and regulators do not have the data to see who is on which side of the gate. They cannot say whether minority-language media have disappeared from automated answers; whether civic agents are producing a common informational baseline or fragmented informational silos; whether the disappearance of investigative reporting from agent outputs reflects market exclusion, alignment policy or both. They are governing an environment they cannot describe.

7.2 The regulated path

On the other path, in 2035, the same citizen asks the same question. The answer they receive may still be concise, personalised and synthesised. But it sits inside an information environment that has been progressively made legible. National media regulators publish annual pluralism assessments that integrate ownership, funding, audience and editorial-independence data with Automated Information Infrastructure data on interfaces, licensing inclusion, source attribution, model and cloud dependency, and alignment patterns. The investigative outlet that first uncovered the policy issue appears in the answer because it is part of a publisher-eligibility framework, accessible to small and local publishers as well as large ones, that ensures the inclusion of public interest journalism in major automated systems. The minority-language newspaper appears because language coverage is a monitored indicator and gaps trigger remedial action. The expert disagreement appears because alignment policies have been tested, in their treatment of contested civic questions, against pluralism criteria that the regulator can articulate and enforce.

The licensing market still exists, but it is no longer the only route into automated information systems. Tier 1 of the framework (interface dominance, source attribution, licensing inclusion or exclusion, crawler access, publisher traffic effects) is now routinely measured. Tier 2 (ownership, strategic partnerships, compute and model dependencies, public-sector procurement) is collected through coordinated work between media regulators, competition authorities and procurement bodies. Tier 3 (full training data, internal alignment, proprietary evaluation methods) is still imperfectly visible, but disclosure regimes built progressively through the 2026–2030 period have made parts of it auditable for the first time. Citizens encounter automated answers that include attribution, that reflect the diversity of the available evidence, and that are produced by systems whose major dependencies are publicly known.

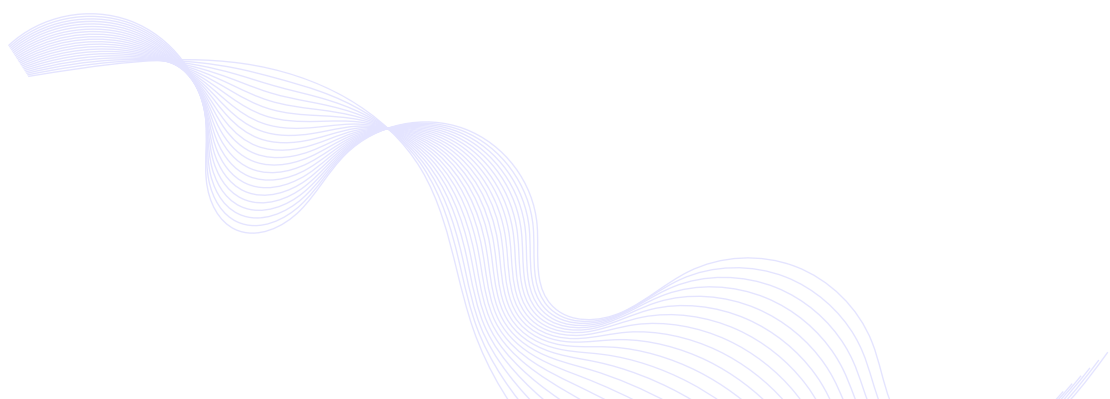
None of this required the regulator to inspect every model or certify every answer. It required the data to exist, and to be available to actors with the authority and the institutional independence to act on it. Journalism remained essential and continued to do its work, with a regulatory infrastructure that no longer treated automated mediation as a separate problem from media pluralism.

7.3 The choice

The two paths diverge from the same starting point: 2026, the moment at which this study has been written. The instruments to take the second path exist, in incomplete form, today. The European Media Freedom Act, the Digital Services Act, the Digital Markets Act, the AI Act, the established competition and data-protection authorities, the academic and civil-society infrastructure, the trade journalism that has already begun reporting on the licensing market and on AI-driven traffic decline: each of these is a piece of the second path. None of them, on their own, is sufficient. Together, with a framework that integrates them and the institutional will to apply it, they are the most complete starting infrastructure that pluralism policy has ever had at the moment a major shift in the information environment was occurring.

What is needed is not new in kind but rather an extension of an existing tradition into the environment now actually structuring how citizens come to know what they know. Classical pluralism indicators should continue to matter. Ownership, funding, editorial independence and reach will remain necessary. But without Automated Information Infrastructure mapping in addition, in a matter of years if not months regulators will not know whether automated systems are amplifying or eroding pluralism; whether licensing markets have created an epistemic elite; whether minority-language media have disappeared from automated answers; whether civic agents are producing a common informational baseline or personalised informational silos. The choice is institutional. Regulators can continue to treat automated systems as a technical supplement to existing media markets, or they can recognise the AIO as a new layer of public sphere infrastructure and adapt their instruments accordingly. The first path leaves pluralism research behind the environment it is supposed to describe. The second extends the tradition now when extension is most needed.

In 2035, the citizen who asks a civic agent about a government decision will be living in one of the two information environments described above, or in a mixture of the two. Which it is will not be determined by any single decision taken by any single actor. It will be determined by the cumulative pattern of choices made, in regulators' offices, publishers' boardrooms, research consortia, civil society organisations, education ministries and newsrooms, between 2026 and the end of this decade. None of those choices is foreclosed. Most of them are still, in 2026, fully open.



References

- Abiri, G. (2024). Generative AI as digital media. *Harvard Journal of Sports & Entertainment Law*, 15(2). <https://ssrn.com/abstract=4878339>
- Baker, C. E. (2006). *Media concentration and democracy: Why ownership matters*. Cambridge University Press.
- Brantner, C., Karlsson, M., & Kuai, J. (2025). Sourcing behavior and the role of news media in AI-powered search engines in the digital media ecosystem: Comparing political news retrieval across five languages. *Telecommunications Policy*, 49(5), 102952. <https://doi.org/10.1016/j.telpol.2025.102952>
- Carlini, R. (2025). Market power and opinion power: Assessing the risks for media pluralism and editorial independence deriving from ownership concentration. *Rivista Italiana di Informatica e Diritto*, 7(1), 75-86
- Chapekis, A., & Lieb, A. (2025, July 22). Google users are less likely to click on links when an AI summary appears in the results. Pew Research Center. <https://www.pewresearch.org/short-reads/2025/07/22/google-users-are-less-likely-to-click-on-links-when-an-ai-summary-appears-in-the-results/>
- Cloudflare. (2025, July 1). Cloudflare just changed how AI crawlers scrape the Internet at large; permission-based approach makes way for a new business model. <https://www.cloudflare.com/press-releases/2025/cloudflare-just-changed-how-ai-crawlers-scrape-the-internet-at-large/>
- Colom, A., Murray, E., & Poblet, M. (forthcoming 2026). *Media, Democracy and Generative AI (I): A Critical Juncture*. A literature review of the current context and ways forward. The Data Tank.
- Doyle, G. (2002). *Media ownership: The economics and politics of convergence and concentration in the UK and European media*. Sage.
- Dragomir, M. (2025). State Media Monitor Global Dataset 2025. Media and Journalism Research Center (MJRC). Zenodo. <https://doi.org/10.5281/zenodo.17219015>
- Dragomir, M. (2019). *Media capture in Europe*. Media and Journalism Research Center / Center for Media, Data and Society.
- European Audiovisual Observatory. (2026). MAVISE: Database on audiovisual services and their jurisdiction in Europe. <https://mavise.obs.coe.int/>
- European Parliament and Council of the European Union. (2022). Regulation (EU) 2022/1925 on contestable and fair markets in the digital sector (Digital Markets Act). Official Journal of the European Union.

European Parliament and Council of the European Union. (2022). Regulation (EU) 2022/2065 on a single market for digital services (Digital Services Act). Official Journal of the European Union.

European Parliament and Council of the European Union. (2024). Regulation (EU) 2024/1083 establishing a common framework for media services in the internal market (European Media Freedom Act). Official Journal of the European Union.

European Parliament and Council of the European Union. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (AI Act). Official Journal of the European Union.

Gillespie, T. (2014). The relevance of algorithms. In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds.), *Media technologies: Essays on communication, materiality, and society* (pp. 167–194). MIT Press.

https://www.microsoft.com/en-us/research/wp-content/uploads/2014/01/Gillespie_2014_The-Relevance-of-Algorithms.pdf

Helberger, N., Karppinen, K., & D'Acunto, L. (2018). Exposure diversity as a design principle for recommender systems. *Information, Communication & Society*, 21(2), 191–207. <https://doi.org/10.1080/1369118X.2016.1271900>

Karppinen, K. (2013). *Rethinking media pluralism*. Fordham University Press.

Keller, P. (2025). *Beyond AI & copyright*. Open Future. <https://openfuture.eu/publication/beyond-ai-and-copyright/>

Mansell, R., Durach, F., Kettemann, M. C., Lenoir, T., Procter, R., Tripathi, G. P., & Tucker, E. (2025). *Information ecosystems and troubled democracy: The state of knowledge on news media, AI, and data governance*. Nordicom, University of Gothenburg. <https://doi.org/10.48335/9789189864238>

Media Pluralism Monitor. (2025). *Media Pluralism Monitor*. Centre for Media Pluralism and Media Freedom, European University Institute. <https://cmpf.eui.eu/projects-cmpf/media-pluralism-monitor/>

Microsoft. (2026, April 27). *The next phase of the Microsoft–OpenAI partnership*. <https://blogs.microsoft.com/blog/2026/04/27/the-next-phase-of-the-microsoft-openai-partnership/>

Molitorisz, S., & Attard, M. (2024). Is Australia's News Media Bargaining Code a model for saving journalism? *Global Perspectives*, 5(1), 126792.

Narayanan Venkit, P., Mahajan, D., Vucetic, S., & Jurgens, D. (2025). Search engines in the AI era: Large language models and the future of search. *Proceedings of the ACM on Human-Computer Interaction*, 9(CSCW1). <https://doi.org/10.1145/3715275.3732089>

Newman, N. (2026). Journalism, media, and technology trends and predictions 2026. Reuters Institute for the Study of Journalism. <https://doi.org/10.60625/RISJ-PS1D-NP11>

News Corp. (2024, May 22). News Corp and OpenAI sign landmark multi-year global partnership. <https://newscorp.com/2024/05/22/news-corp-and-openai-sign-landmark-multi-year-global-partnership/>

Nielsen, R. K., & Ganter, S. A. (2022). The power of platforms: Shaping media and society. Oxford University Press. <https://doi.org/10.1093/oso/9780190908850.001.0001>

Noam, E. M. (2009). Media ownership and concentration in America. Oxford University Press.

Noam, E. M. (Ed.). (2016). Who owns the world's media? Media concentration and ownership around the world. Oxford University Press.

Ofcom. (2025). The era of answer engines: Generative AI's impact on search experiences and online safety. Ofcom. <https://www.ofcom.org.uk/internet-based-services/technology/the-era-of-answer-engines-generative-ais-impact-on-search-experiences-and-online-safety>

OpenAI. (2023, December 13). Partnership with Axel Springer to deepen beneficial use of AI in journalism. <https://openai.com/index/axel-springer-partnership/>

OpenAI. (2024, May 22). A landmark multi-year global partnership with News Corp. <https://openai.com/index/news-corp-and-openai-sign-landmark-multi-year-global-partnership/>

OpenAI. (2026, April 27). The next phase of the Microsoft–OpenAI partnership. <https://openai.com/index/next-phase-of-microsoft-partnership/>

Pattison, C., Ricks, V., & Wihbey, J. (2025, August 11). How AI-driven search may reshape democracy, economics, and human agency. Tech Policy Press.

Pickard, V. (2020). Democracy without journalism? Confronting the misinformation society. Oxford University Press. <https://doi.org/10.1093/oso/9780190946753.001.0001>

Radsch, C. (2024). What is the value of journalism to AI? Center for Journalism & Liberty.

Radsch, C., & Montoya, K. (2026). Same gatekeepers, new tollbooths: Mapping the AI content licensing market. Center for Journalism & Liberty / Open Markets Institute.

Reuters Institute for the Study of Journalism. (2026, January 13). News publishers expect search traffic to fall by more than 40% in the next three years, new RISJ report finds. University of Oxford.

Schiffrin, A., Kostic, B., Nenadic, I., Haas, J., Helberger, N., & van Drunen, M. Z. (2025). Safeguarding media freedom in the age of Big Tech platforms and AI. OSCE Representative on Freedom of the Media.

Schneiders, P., & Stark, B. (2025). Ensuring news quality in platformized news ecosystems: Shortcomings and recommendations for an epistemic governance. *Media and Communication*, 13.

Seipp, T. J., Helberger, N., de Vreese, C., & Ausloos, J. (2024). Between the cracks: Blind spots in regulating media concentration and platform dependence in the EU. *Internet Policy Review*, 13(4). 1-26. <https://doi.org/10.14763/2024.4.1813>

Simon, F. M., Nielsen, R. K., & Fletcher, R. (2025). Generative AI and news report 2025: How people think about AI's role in journalism and society. Reuters Institute for the Study of Journalism. <https://doi.org/10.60625/RISJ-5BJV-YT69>

Sjøvaag, H. (2024). The business of news in the AI economy. *AI Magazine*, 45(2), 246–255. <https://doi.org/10.1002/aaai.12172>

Van den Bulck, H., Kulig, M., Hyzen, A., Puppis, M., & Paulussen, S. (2025). Epistemic welfare and public service media's algorithmic recommender systems: A theoretical framework, operationalization and relevance for governance. *European Journal of Communication*, 40(6), 598-618. <https://doi.org/10.1177/02673231251385761>

van Dijck, J., Poell, T., & de Waal, M. (2018). *The platform society: Public values in a connective world.* Oxford University Press. <https://doi.org/10.1093/oso/9780190889760.001.0001>

MJRC

MEDIA & JOURNALISM
RESEARCH CENTER

Media and Journalism Research Center

Legal address

Tartu mnt 67/1-13b, 10115,
Tallinn, Harju Maakond, Estonia

Postal address

6 South Molton St, London,
W1K 5QF, United Kingdom

Academic affiliation

Universidade de Santiago de Compostela (USC)
Colexio de San Xerome, Praza do Obradoiro s/n,
CP 15782 de Santiago de Compostela.

Contact

www.journalismresearch.org
mjrc@journalismresearch.org



Artificial Intelligence (AI) Disclosure Statement

No AI tools were used in this research project for data collection. No AI tool was used in drafting or shaping the analytical content of this study. Grammarly was used to correct grammar during the editing phase. ChatGPT was used for a final pre-publication proofreading and formatting check

doi: 10.5281/zenodo.20322621

File integrity (SHA-256):

3DD8CAF7AF37E5975EA9C92416AC36CE70336BEA81835883BC1
8E539D0FA72B7