

DIGIPOL Researcher Toolkit

DIGIPOL – Deliverable D8

Zach Bastick

European University Institute (EUI), Florence

1 August 2025

Funding statement

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



Funded by the
European Union

Toolkit Contents

Introduction..... 3

Alone and in the dark: Sensemaking through online platforms during the COVID-19 pandemic..... 4

Governance of Visibility: How Social Media Interfaces Allocate Power..... 13

Beyond Platform Transparency: Ten Policy Recommendations for Digital Public-Interest Research.... 21

Introduction

This researcher toolkit brings together a selection of key texts produced by the DIGIPOL project. It is designed as a curated compilation of articles that reflect the project's broader intellectual development: from questions of digital sensemaking and platformed mediation, to the governance of visibility, and finally to the policy implications of public-interest research in platformed environments.

Taken together, these texts show how the project evolved from an interest in platform transparency toward a more developed concern with the conditions under which digital environments can be independently observed, interpreted, and governed in the public interest.

The articles included here are complementary in scope and method. Some are more conceptual, seeking to clarify how platforms shape knowledge, attention, and observation; others are more directly oriented toward research and policy, drawing out practical implications for scholars, regulators, and democratic oversight. Read together, they offer a set of resources for researchers working on platforms, digital politics, and the changing infrastructures of online visibility.

This toolkit is therefore intended as a usable research resource: a consolidated entry point into some of the main ideas, arguments, and outputs developed through DIGIPOL. It is offered in the hope that it will support further reflection and inquiry into how platformed information environments can be studied critically, and how the conditions for meaningful public-interest observability might be strengthened.

Alone and in the dark: Sensemaking through online platforms during the COVID-19 pandemic

Zach Bastick

"In lockdown... alone.
A COVID host to be a ghost?
Surfing endlessly from home.
Not me. I'll get free."

19 words on COVID-19 (Bastick, this essay)

Keywords: *Algorithmic governance, content moderation, sensemaking, COVID-19, filter bubbles, misinformation, online platforms*

Introduction

In the years leading up to 1920, the French surrealists devised a method of collectively compiling works of poetry and drawing. Typically sitting in a Parisian café, artists took turns to contribute a few words to a poem, or a few strokes to an illustration. The final product would incrementally take form as it moved from person to person. This method of communally producing creative works – known as exquisite corpse – has as an attribute that it isolates the individual in a collective process. It is suggestive of systems: To the engineer who revels in the predictability of Boolean logic, it highlights the mechanics of information systems, with their sequential inputs, processes, and outputs. To the philosopher who revels in stories of unpredictability, such as of hurricanes being caused by the flap of a butterfly wing on the other side of the globe, it highlights how seemingly small and naïve contributions can produce large and surprising effects. It also highlights how individuals have come together online to make sense of the COVID-19 pandemic - with one person's posts and retweets informing another person's posts and retweets, and so on, until a communal but often temporary sense of reality emerges.

The poem at the beginning of this chapter was the initial input to an exquisite corpse – a first flap of a wing. I wrote it while in lockdown in Paris, sitting at a computer at home (instead of a café), in an environment replicated by dozens of other first world cities around the world: connected through the internet, guarded by the confines of a small apartment, and threatened by logistical and physical structures that breached those walls. It was prompted by a prompt of the *Massive/Micro* experiment to produce a 19-word script on sense-making between humans, machines, and the planet during the COVID-19 pandemic. Lynnette Widder at Columbia University, also while in lockdown (but on the other side of the Atlantic), produced a series of illustrations of this poem, shown below. Elizabeth Whitney from the City University of New York and Heidi A. Campbell from Texas A&M University also contributed. This produced a moving image that sought to reflect our experiences of the pandemic as individual social scientists and media scholars. This was a minute in lockdown: largely inconsequential to the outside world, but a snapshot of a shared journey. It was also a deliberate manifestation of the massive sensemaking process that concurrently occurred organically and digitally all over the world.



Fig 1: Line-by-line illustration of the poem by Lynette Widder (with permission from Lynette Widder)

In reality, of course, the processes of sensemaking in society are more complex than can be described by the exquisite corpse: they involve individuals but also communities, who interpret the world through very different lenses, both simultaneously and over time, and produce a multitude of worldviews, which translate to parallel and often competing realities. In addition, while lockdowns encouraged a migration to digital spaces (Bao et al., 2020; Nabity-Grover et al., 2020), praise of the internet as a ‘pandemic safe’ information source is tempered by concerns about authenticity and trust. The extent of COVID-19 misinformation has been described as an ‘infodemic’ by public health experts and combating this has become a priority for transnational organizations and national governments (Radu, 2020). The politicization of the pandemic has incentivized disinformation, especially within the context of the 2020 US presidential election (Faris et al., 2020; Mian & Khan, 2020). In parallel, there is increasing discussion about platform governance, filter bubbles, echo chambers, and the effects of algorithms – all of which artificially mediate the information that we access online (Gorwa, 2019; O’Neil, 2016).

During lockdown and times of social distancing, online spaces provided various answers to the questions raised by COVID-19: How did the virus come to be? How is it spreading? What is the true extent of the pandemic? Will the pandemic die down naturally? What are the symptoms and fatality rate? How can one protect oneself from infection? Conflicting answers to those questions in turn raise auxiliary questions, including some of the most fundamental of our epistemological struggles: ‘What is real?’ and ‘What is true?’. Our dependence on the internet in a moment of crisis raises the issue of how

we use the internet to make sense of COVID-19. More specifically, it raises the question of how we experience sensemaking online, and how we individually balance the digital affordances and pitfalls of the modern digital environment. In this chapter, I autoethnographically draw on my online experience during the early stages of COVID-19. I ask you, dear reader, to consider the following question: How does your awareness of the dynamics of online spaces affect your journey to make sense of the world?

Sensemaking as an experience

Sensemaking is the purposeful process by which one moves from not understanding something to understanding it. A long history of psychology literature proposes that we group together facts, attributes, and events. These groups feed mental models that help us to understand the world. In exploring how intelligence develops in children, for example, Jean Piaget purported that children are born with innate cognitive schemata that they use to interpret the world. Piaget understood a schema to mean a "cohesive, repeatable action sequence possessing component actions that are tightly interconnected and governed by a core meaning" (Piaget 1952, p. 7). By observing the behavior of his three young children, he concluded that these schemata develop organically through a process of assimilation and accommodation. That is, according to Piaget, we use schemata to interpret the information that we encounter. We then assimilate this information into our schemata, which we then apply to bolster our interpretation of future information; or, if the information does not fit the schemata, we modify our schemata.

This notion of schemata – or more generally, of *knowledge structures* – has been rearticulated in different contexts as stories, rules, images, common attributes, scripts, and templates. The need to operationalize sensemaking, such in the design of Artificial Intelligence, has led these various notions to be aggregated to the concept of a *frame* (Klein et al., 2007). Karl E. Weick (1995), who introduced the notion of sensemaking to organizational studies, referred to frames as the 'structure of context'. He viewed frames as compartments into which people place and organize the data they encounter in the world. A decade later, the relationship between frames and data was conceptualized more procedurally by Klein et al (2006) in the data-frame model. They viewed this relationship between frame and data as symbiotic and cyclical: frames help us to interpret data, but data also modifies frames. New data might enable us to draw new connections within a frame, or we might encounter data that does not fit the frame and necessitates the frame to be refined or replaced. From this, various processes of sensemaking can be described.

A linear conception of sensemaking is easy to understand: One encounters a Twitter post, YouTube video, or online news article, and interprets it according to a frame, which is in turn modified by the new information. As noted in the introduction, however, neither sensemaking nor autoethnography is necessarily linear nor encapsulated; and the experience of sensemaking is richer than that which a procedural and positivist approach alone can encapsulate. Weick (1995) himself has stressed additional aspects of sensemaking: that it is grounded in identity, that it is enactive of social environments, and that it is driven by plausibility rather than accuracy. The influential theory of sensemaking proposed by Brenda Dervin (1992; 2015), while originally focusing on cognitive and procedural information behaviors, such as seeking and processing information, has gradually evolved to emphasize verbing and structural powers or 'energies' which mediate sensemaking (Savolainen, 2006).

In line with this expanded notion of sensemaking, Annette Markham has argued that understanding the digital world requires that we become cognizant of our frames beyond what is traditionally expected in scientific culture. She criticizes academia for presenting knowledge as a final, often publishable product (rather than a continual exchange), and data as an object isolated from inquiry (rather than a source that shifts through the experience of inquiry, the latter of which produces data in itself). Markham recommends emphasizing "meaning versus method; engagement versus objectivity;

interpretation versus findings; argument versus explanation” (Markham, 2013, section 4.2, para. 9) as a more realistic and ethical means of making sense of the world.

Autoethnography, similarly, highlights the salience of the researcher (their experience, identity, and perspective) in sensemaking, and brings to the forefront the limitations and affordances of personal experience (Holman Jones & Harris 2018; Adams et al., 2015). The complexity of sensemaking – the interrelatedness between and within the macro and micro of data and interpretation – became a challenge during the coronavirus pandemic. Epidemiologists remarked on the short time that the virus took to spread from its first identified source in Wuhan, China to every continent on earth (Chinazzi et al., 2020; Khachfe et al., 2020). While epidemiological models highlight the interconnectedness of human beings, they fail to capture many of the nuances of the pandemic (Enserink & Kupferschmidt, 2020). Indeed, the data produced during the pandemic is vast, and transcends the disciplinary confines of medicine, sociology, and political science. Perspective, too, is important, and Angeli and Montefusco (2020) stress that the interpretation of data depends in part on the political context, public health, and the time during which it is collected. To find sense in this agglomeration, they advocate a complex systems approach to pandemic sensemaking. As Joachim Sturmberg and Carmel Martin note in advocating a similar perspective, the “pandemic reveals that everything is connected to everything else” and this requires us to re-analyze our systems of sensemaking (Sturmberg & Martin 2020, p.1).

Despite the need for tangible policy proposals and empirical evidence, it should be remembered that sensemaking is a profoundly human activity. Sensemaking occurs at the collective, civilizational, and global levels, but also at the personal level. It involves the processing of information and frames, but also the externalities of this process: our hesitations, social bonds, dialogue with others, anxieties, and behaviors, to name but a few. Personal context mediates the process of sensemaking, but also situates it within our lives. That is, we must remember that sensemaking is an *experience* as well as a narrower *process*. This enables us to shed light not only on inquiry, but also on our relationship with inquiry and our relationship with the objects of our inquiry. Hence, with people increasingly relying heavily on the internet to make sense of the pandemic, it is important to examine our experiences of online sensemaking.

Autonomy and shared space

My first recollection of experiencing sensemaking through online environments came from building a website in 1997 which hosted *dancing baby* videos. These videos were “all the rage on the net” (McNamara 1997, p. 39), as Network World magazine reported at the time, with internet users delighted at the sight of an animated baby dancing. These videos also illustrated that digital technology reconfigured reality (where else could you see a baby dancing like an adult?). A piece of Software called 3D Studio Max enabled anyone with basic animation skills to create a new dancing baby. As a result, a delightfully haphazard series of videos emerged online, becoming what might be regarded as the first graphical internet.

There was a certain thrill to watching the numbers going up on my website’s traffic counter. I recall trying to imagine all these people in a room – first to grasp the sheer quantity of real people behind the hit counter, and then to imagine what those people might look like and how they might interact. In this vein, I placed a guestbook on the website, which attracted comments that ranged from praise to threats. Perhaps the ability to comment anonymously unleashed facets of the human experience that society repressed. Indeed, it felt as though on this early web, individuals could disembodied their offline personalities and explore new facets of the psyche - as *The New Yorker* had phrased it four years earlier, “on the internet, nobody knows you’re a dog” (Steiner, 1993). This enabled a liberal and almost anarchic space in which everyone could be a ruler of themselves, precisely because their modems had set them free.

Decades later, a threat posted on this guestbook still sticks with me. It spared no detail in the vitriolic, gruesome, and vivid description of torture that would be performed on me. Being 11 years old, I recognized their words by their purely literal meaning, but had never experienced their forging into forceful language. I was left with a concatenation of questions that my current frames could not answer: To what extent were these threats contemplated, and to what extent were they instinctive lashes at a keyboard? What place did they have on my amateur dancing baby website (and should I delete them)? Who was the author of these threats and what circumstances motivate them to bring such darkness into my world? Would they have typed the same sentences knowing that I was only a young boy? I wondered how their experience of the world differed from mine, and how the contexts in which they lived could feel so alien to me. That is, I wondered about *their* frames and the limitations of *my own* frames.

This was a visceral, animalistic, and unforgiving experience in online sensemaking. It made me recognize that the internet empowers sensemaking by placing the heterogeneous experiences of humanity in a shared online space. It showed me, and made me feel, that the internet affords autonomy: autonomy to break offline boundaries, autonomy to explore the contexts of others, autonomy to be creator and participant, and autonomy to be yourself. The online world has now changed – as have the ways that we understand ‘autonomy’ and ‘shared space’ online - but reflecting on our online experiences can highlight that the experience of online sensemaking is richer than the process of online sensemaking.

The evisceration of online sensemaking during COVID-19

Today, it might seem quaint to associate ‘autonomy’ and ‘shared space’ with online sensemaking. It suggests the outmoded optimism of early digital communities, and hints at the romanticism of John Perry Barlow’s Declaration of the Independence of Cyberspace (“property, expression, identity, movement, and context do not apply to us. They are all based on matter, and there is no matter here” [Barlow 1996, para. 9]). For many, the 2018 Cambridge Analytica Scandal was a wakeup call to changes in our online environments. Media scholars now decry the many emerging dangers of online spaces, which need not be repeated here to make the point that we are becoming increasingly reflective of our online environments. Given this, it is pertinent to ask how this reflexivity affects the ways in which we make sense of the world through online spaces.

In spite of the dampening optimism of media scholars, the internet was an important source of information on the pandemic. Social media relayed the initial attempts of the Chinese government to suppress news of the COVID-19 outbreak, including the government admonishing Li Wenliang, the physician who first alerted the world to the COVID outbreak on WeChat (Green, 2020). Amid fears of a government cover-up, researchers around the world dissected carefully worded statements by the Wuhan Municipal Health Commission and used the web to speculate on the appearance of a new coronavirus, SARS-CoV-2 (Gralinski & Menachery, 2020). Social media enables researchers to share emergent findings and healthcare whistleblowers to expose overburdened healthcare facilities. I came to fully realize the urgency of social distancing; not from a government campaign or epidemiological study, but from a post on the social journalism platform Medium that had been shared on Twitter. Within nine days of its publication on March 20, 2020, the post had gained over 40 million views and been translated to over 40 languages (Pueyo, 2020). Just as public authorities strived to grasp the rapidly unfolding situation, social media allowed millions around the world to collaboratively make sense of COVID-19. Likewise, I relied on the internet to make sense of the early days of the pandemic. The internet also provided me with a ‘digital’ form of stability amid the ‘biological’ turmoil of the pandemic. I recount the experience below:

I am in Paris, France, on the eve of the first national lockdown. On March 16, 2020, at 8pm, I turn on the television to watch French president Emmanuel Macron announce the lockdown, beginning at noon the next day. The French lockdown entails downloading an authorization form or writing out a sworn testimony to leave one's home, under penalty of being fined by police. It feels like the world is moving; yet from one day to the next, the city turns silent as the streets empty of their traffic. At this moment, smartphones, tablets, and computers, offer an obvious way to remain connected to the world while being separated from it. For my part, I make sense of the pandemic through online news, social media, and COVID-19 data portals. Online news is my primary source of authoritative information on the pandemic, and I use social media to gain more granularity on public discourse. I also find myself reloading, multiple times a day, the Johns Hopkins Coronavirus Dashboard (Dong et al., 2020). For the next months, that dashboard is often the first website that I check in the morning and the last that I check at night.

It has long been noted that the algorithms that curate newsfeeds and online news aim to not only return relevant data but to determine the criteria of what is relevant (Das et al., 2007). Evidence of news aggregators producing filter bubbles or reinforcing political stances is mixed (Bruns, 2019; Haim et al., 2018; Le et al., 2019). However, the politicization of the pandemic increased incentives to control public information – and that elevated the role of trust in using algorithmically mediated spaces for information provision. This becomes particularly real when considering the 'algorithm' as a metonym that hides complex power structures. These power structures involve real people making real decisions, and so reflect the values and interests of those who own, control, and exploit these algorithms. They also involve the machine uncertainty of artificial intelligence outcomes. The question that arises here, then, is not how the limitations of the internet affect online sensemaking (although this is a much asked and important question). Rather, the question that arises is how acknowledging these influences over online spaces affects the experience of online sensemaking.

Perhaps like yours, my online sensemaking journey began with the desire to answer immediate questions (such as 'How did the virus come to be?' and 'What is the true extent of the pandemic?'). You may have experienced moments of insight while seeking out answers to these questions. For example, my social media feeds emphasized the benefits of national lockdowns, the importance of personal protective equipment, and the failed responses of national governments worldwide – but I came to realize that social media was isolating me from alternative narratives:

One day, I click on a YouTube video on a topic outside of my typical interests; then I click on a recommended video; and then I break the filter bubble. The comments on the video universally oppose lockdowns, decry the use of personal protective equipment, and praise the Trump administration's handling of the pandemic. It feels as if I have entered a parallel universe – a feeling cemented by one comment noting that the clear agreement in the comment section was evidence of establishment media silencing 'real people'. Have I escaped my filter bubble to view reality? Or have I entered a new filter bubble, with its alternative mirage of reality? It strikes me like a bat over the head at that moment, that I have no obvious way of determining the effects of the invisible decisions of platforms on my frames, and no obvious way of situating these frames in relation to anyone else's. In moments like these, you feel alone and subjugated, rather than connected and free; and you have no clearer understanding of the world than when you began.

Reflexively analyzing our personal experiences in online sensemaking can shift our questions. For me, my immediate questions were supplanted by deeper epistemological and ontological questions, such

as: How can you feel so isolated on the internet, a network that connects you to millions? How do you determine the extent to which your freedoms are limited by obscure digital influences? If you have no sense of shared space or objectivity, how can you determine the extremity of your own views? More fundamentally: To what extent can reality be shared? What does it mean to exist in a virtual world where your frames, and consequently thoughts and behaviors, are shaped by a relatively small group of people who govern online platforms? These were raised by an awareness that I no longer experienced a sense of shared space and autonomy online. It illustrates how becoming aware of experience of online sensemaking can modify our lines of inquiry and change the assumptions of our online sensemaking processes.

Conclusion

It is paradoxical to feel isolated in lockdown while connected to millions online; and, in turn, to be connected to masses of information but disconnected from reality. These incongruities evidence the complexity of online sensemaking. In this chapter, I applied a critical autoethnography (Harris & Holman Jones 2019) to explore online sensemaking during the pandemic. This aimed to highlight that while we are increasingly aware that we use online platforms to make sense of the world, and of the dynamics of these platforms, a further step is to recognize that our awareness of these platforms also affects our sensemaking experiences.

More generally, being reflexive – in focusing on the full human *experience* of sensemaking rather than the narrower *process* of sensemaking – can encourage us to explore fundamental questions about sensemaking in the digital world. These questions can seem daunting at first ('if I don't know what is real, where do I go from here?'), but they can also be liberating. They highlight that sensemaking is more than a process – it is the experience of life.

As Daniel Harris, Mary Elizabeth Luka, and Annette Markham encouraged me to write about how we make sense of ourselves, others, and the world, and asked how the pandemic helped to make sense of the relationship between humans, machines, and the planet, I struggled with the contradiction being connected to the internet while also feeling alone and in the dark. This struggle is displayed in the poem at the beginning of the chapter and in the experience of online sensemaking.

You, too, might find that your awareness of digital platforms impacts the ways you make sense of the world. So, I invite you to reflect about your own experience of sensemaking online, especially during moments of heightened uncertainty: First, how does the internet affect your sensemaking processes? Then, how does being aware of this affect your sensemaking experience? Finally, what does this tell you about yourself, others, technology, and the world? If this inquiry raises new questions that appear destabilizing, you may find the act of reflection itself liberating. Hence, I encourage you to consider grappling with these questions – as reflexivity can manifest your role not only as a 'user of the internet', but also as an active and conscious participant of the digital world.

References

- Adams, T. E., Jones, S. H., & Ellis, C. (2015). *Autoethnography: Understanding Qualitative Research*. London: Oxford University Press.
- Angeli, F., and Montefusco, A. (2020). Sensemaking and learning during the Covid-19 pandemic: A complex adaptive systems perspective on policy decision-making. *World Development*, 136, 105106.
- Bao, H., Cao, B., Xiong, Y., and Tang, W. (2020). Digital Media's Role in the COVID-19 Pandemic. *JMIR MHealth and UHealth*, 8(9), e20156.
- Bruns, A. (2019). *Are Filter Bubbles Real?* John Wiley & Sons.
- Chinazzi, M., Davis, J. T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., Piontti, A. P. y, Mu, K., Rossi, L., Sun, K., Viboud, C., Xiong, X., Yu, H., Halloran, M. E., Longini, I. M., and Vespignani, A. (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science*, 368(6489), 395–400. <https://doi.org/10.1126/science.aba9757>
- Das, A. S., Datar, M., Garg, A., and Rajaram, S. (2007). Google news personalization: Scalable online collaborative filtering. *Proceedings of the 16th International Conference on World Wide Web*, 271–280. <https://doi.org/10.1145/1242572.1242610>
- Dervin, B. (1992). From the mind's eye of the user: The sense-making qualitative-quantitative methodology. *Sense-Making Methodology Reader*.
- Dervin, B. (2015). Dervin's Sense-Making Theory. In *Information seeking behavior and technology adoption: Theories and trends* (pp. 59–80). IGI Global.
- Dong, E., Du, H., and Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases*, 20(5), 533–534. [https://doi.org/10.1016/S1473-3099\(20\)30120-1](https://doi.org/10.1016/S1473-3099(20)30120-1)
- Enserink, M., and Kupferschmidt, K. (2020). With COVID-19, modeling takes on life and death importance. *Science*.
- Faris, R., Clark, J., Etling, B., Kaiser, J., Roberts, H., Schmitt, C., Tilton, C., and Benkler, Y. (2020). *Polarization and the Pandemic: American Political Discourse, March – May 2020*. <https://dash.harvard.edu/handle/1/37366128>
- Gorwa, R. (2019). What is platform governance? *Information, Communication & Society*, 22(6), pp. 854–871.
- Gralinski, L. E., and Menachery, V. D. (2020). Return of the Coronavirus: 2019-nCoV. *Viruses*, 12(2), p. 135. <https://doi.org/10.3390/v12020135>
- Green, A. (2020). Li Wenliang. *The Lancet*, 395(10225), 682.
- Haim, M., Graefe, A., and Brosius, H.-B. (2018). Burst of the Filter Bubble? *Digital Journalism*, 6(3), pp. 330–343.
- Harris, A. and Holman Jones, S. (2019). *The Queer Life of Things*. NY: Lexington/ Rowman & Littlefield.
- Holman Jones, S. and Harris, A. (2018). *Queering Autoethnography*. NY: Routledge.
- Khachfe, H. H., Chahrour, M., Sammour, J., Salhab, H., Makki, B. E., and Fares, M. (2020). An epidemiological study on COVID-19: A rapidly spreading disease. *Cureus*, 12(3).
- Klein, G., Moon, B., and Hoffman, R. R. (2006). Making sense of sensemaking 2: A macrocognitive model. *IEEE Intelligent Systems*, 21(5), pp. 88–92.
- Klein, G., Phillips, J. K., Rall, E. L., and Peluso, D. A. (2007). A data-frame theory of sensemaking. *Expertise out of Context: Proceedings of the Sixth International Conference on Naturalistic Decision Making*, pp. 113–155.
- Le, H., Maragh, R., Ekdale, B., High, A., Havens, T., and Shafiq, Z. (2019). Measuring Political Personalization of Google News Search. *The World Wide Web Conference*, pp. 2957–2963. <https://doi.org/10.1145/3308558.3313682>
- Markham, A. N. (2013). Undermining 'data': A critical examination of a core term in scientific inquiry. *First Monday*. <https://doi.org/10.5210/fm.v18i10.4868>
- McNamara, P. (1997, June 16). Baby talk: This twisting tot is all the rage on the 'Net. *Network World*, 14(24), p. 39.

- Mian, A., and Khan, S. (2020). Coronavirus: The spread of misinformation. *BMC Medicine*, 18(1), p. 89. <https://doi.org/10.1186/s12916-020-01556-3>
- Nabity-Grover, T., Cheung, C. M. K., and Thatcher, J. B. (2020). Inside out and outside in: How the COVID-19 pandemic affects self-disclosure on social media. *International Journal of Information Management*, 55, 102188. <https://doi.org/10.1016/j.ijinfomgt.2020.102188>
- O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway Books.
- Piaget, J. (1952). *Origin of Intelligence in the Child*.
- Pueyo, T. (2020, May 28). *Coronavirus: Why You Must Act Now*. Medium. <https://medium.com/@tomaspueyo/coronavirus-act-today-or-people-will-die-f4d3d9cd99ca>
- Radu, R. (2020). Fighting the 'Infodemic': Legal Responses to COVID-19 Disinformation. *Social Media + Society*, 6(3), 2056305120948190. <https://doi.org/10.1177/2056305120948190>
- Savolainen, R. (2006). Information use as gap-bridging: The viewpoint of sense-making methodology. *Journal of the American Society for Information Science and Technology*, 57(8), pp. 1116–1125.
- Steiner, P. (1993). *On the internet nobody knows you're a dog*. *The New Yorker*. 69(20), p. 61.
- Sturmberg, J. P., & Martin, C. M. (2020). COVID-19—how a pandemic reveals that everything is connected to everything else. *Journal of Evaluation in Clinical Practice*.
- Weick, K. E. (1995). *Sensemaking in Organizations*. Thousand Oaks: Sage.

Funding statement



Funded by the
European Union

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

Governance of Visibility: How Social Media Interfaces Allocate Power

Zach Bastick
European University Institute & Harvard University
zach.bastick@eui.eu

May 2025 Working Paper

Abstract

This article develops a theoretical account of digital platforms as infrastructures and advances the concept of governance of visibility - the strategic control over who can observe, measure, and know platformed publics. Bringing infrastructure studies into dialogue with critical political economy and platform governance, this paper argues that interface decisions (such as APIs, research tools, rate limits, and eligibility rules) act as infrastructural choke points that reallocate epistemic power under digital capitalism. Two recent trajectories - X's paywalled API and Meta's retirement of CrowdTangle - illustrate how visibility has been enclosed and monetized, with downstream effects for public-interest research and accountability. The paper then theorizes counter-infrastructures, mapping the distinct characteristics of dependent and independent counter-infrastructures. The discussion specifies design and policy levers, focusing on the EU Digital Services Act (DSA) Article 40 regime on vetted researcher access and the challenge of turning transparency from corporate favor into enforceable right. Findings are drawn from work undertaken during the EU-funded DIGIPOL project.

1. Introduction

In 2023, Twitter (now X) ended research-viable free access to its API and introduced costly tiers and tighter rate limits. Research centers and scholars publicly warned that the change would curtail a variety of public-interest uses - including crisis informatics, misinformation monitoring, and longitudinal audits - and could mark the end of an era for platform transparency (Freelon, 2018; Tromble, 2021). Many also described how pricing and terms made population-scale observation unaffordable for public institutions and initiatives (Newton, 2023). Some scholars argue that the cumulative effect has been to chill independent oversight by narrowing who can measure the platform at scale (Starbird, 2019; Tromble, 2021).

Likewise, in 2024, Meta decommissioned CrowdTangle, a widely used instrument for observing content diffusion on Facebook and Instagram. Although Meta introduced a replacement (the Content Library/API), journalists and watchdog organizations criticized reduced real-time observability and limits on eligibility, especially during election periods (Associated Press, 2024; Tow Center for Digital Journalism, 2024). Open letters urged Meta to maintain CrowdTangle through major election cycles (Tow Center for Digital Journalism, 2024).

This article makes sense of these developments by theorizing social media as infrastructures of digital capitalism. Rather than analyzing only ranking or moderation decisions, we foreground interfaces - APIs, research tools, documentation, rate limits, and eligibility rules - as infrastructural choke points through which platforms enact governance of visibility. We define 'governance of visibility' as the

strategic control over the means of observability: who can see, at what granularity and latency, and under what legal and financial terms.

The findings from this paper are drawn from the «Digital Observatory of Political Discourse» (DIGPOL), an EU funded project seeking to provide transparency into speech controlled by highly used corporate platforms. The paper theorizes social platforms as infrastructures within digital capitalism, explores the mechanisms through which interfaces govern visibility and constrain public oversight, and examines the rise of counter-infrastructure as it exists today. It then situates this within recent EU legislative developments.

2. Theoretical framework: Platforms as infrastructures in digital capitalism

Infrastructure studies has conceptualized infrastructures as embedded, standardized, and often invisible systems that become visible at breakdown (Star, 1999). They are relational backbones that enable action across sites and scales; their politics are inscribed in standards, gateways, and dependencies. Reading social media as infrastructure foregrounds upstream control points - identity systems, schemas, developer policies, and interfaces - that are usually backgrounded when we only analyze feeds and moderation. Plantin and colleagues capture a contemporary shift as a dual movement: the infrastructuralization of platforms and the platformization of infrastructures (Plantin et al., 2018). Large platforms (e.g., Google, Facebook) acquire infrastructural characteristics - ubiquity, reliability, standard-setting - while traditional infrastructures adopt platform logics. The infrastructural lens matters because infrastructural status historically carried public obligations (interconnection, nondiscrimination, continuity, legibility). When private platforms perform infrastructural roles, interface decisions become matters of public concern (Plantin et al., 2018; Star, 1999).

In parallel, critical political economy explains why infrastructural control over interfaces is central to platform power. Srnicek (2017) situates platforms as intermediaries that extract data and generate network effects, turning ecosystem control into rents. Under this logic, programmability is valuable when it scales the ecosystem, but enclosure is attractive when it preserves proprietary advantage and turns access itself into a revenue line. Zuboff (2019) emphasizes asymmetric visibility: the unilateral capture of behavioral data behind corporate opacity. Platforms build one-way mirrors - near-omniscience over users and publics, themselves having minimal legibility about internal operations - so that prediction and control can be monetized. The enclosure of visibility (what outsiders can know) is thus not peripheral; it is constitutive of the accumulation model. Couldry and Mejias (2019) extend the critique historically: contemporary extraction reorganizes social life so that human experience becomes a continuous input to capital - an appropriation of life through data relations. In short, decisions to close or price APIs, retire tools, or vet researchers are not solely technical or policy decisions, but they are political-economic acts: they privatize the means of observation, protecting accumulation by limiting counter-knowledge (Couldry & Mejias, 2019; Srnicek, 2017; Zuboff, 2019).

Platform governance scholarship shows how firms curate content visibility (ranking, moderation) while presenting themselves as neutral “platforms” (Gillespie, 2010; Gorwa, 2019). We extend this to institutional visibility: who is permitted to see the platform itself - its data flows, interventions, and systemic risks. Pasquale (2015) warns that formal transparency can be performative and insufficient; complex systems can remain unintelligible without institutional arrangements that ensure independence, completeness, and contestability. Ananny and Crawford (2018) similarly argue that accountability is a sociotechnical process, not a property of a dataset or code dump. These insights motivate a conceptual shift from the generic “transparency” often cited in the literature to a governance of visibility.

3. The need for a governance approach

Between February and May 2023, Twitter/X ended research-viable free access, tightened endpoints and rate limits, and rolled out paid tiers whose enterprise pricing exceeded many academic budgets (Newton, 2023). Likewise, in August 2024 Meta retired CrowdTangle, replacing it with the Content Library/API. Journalists, watchdog organizations, and officials criticized the timing (election cycles) and scope (eligibility limits; reduced real-time features), and coalitions urged Meta to maintain CrowdTangle through 2024-2025 (Associated Press, 2024; Tow Center for Digital Journalism, 2024). Meta's replacement may improve access for vetted researchers but curtails a public-facing transparency instrument used by newsrooms and civil society.

Scholars cited impacts on crisis response, misinformation monitoring, and longitudinal analyses (Freelon, 2018; Tromble, 2021). Some scholars argue that the cumulative effect has been to chill independent oversight by making population-scale observability contractually constrained or unaffordable (Starbird, 2019; Tromble, 2021). From a framework of governance of visibility, these developments recalibrate the visibility regime through pricing, throughput controls, and legal constraints. Small-N audits remain possible, although mediated by algorithmic obscurity, but large-scale, real-time analysis becomes a paid privilege governed by terms the platform can change unilaterally. The net result is a redistribution of epistemic power toward the platform and its chosen partners. From our framework, tool deprecation resets the measurement instrument and moves visibility into a more curated channel. Even where replacements exist, latency, granularity, and eligibility are re-specified by the firm, shifting oversight from open instrumentation to a gated portal.

There are multiple modalities through which platforms govern visibility. These including pricing and rate-limit regimes, such as volume and throughput caps, which transform observability into a commodity, selectively affordable for well-resourced actors (Newton, 2023); Qualification and vetting, notably through access regimes delineate eligible observers, often with publication constraints; tool deprecation and substitution, which as in CrowdTangle, collapses measurement ecologies; and through terms of service modifications, which undermine reproducibility and add legal risk to independent data collection (Freelon, 2018; Tromble, 2021).

Much public debate demands 'transparency' in online information ecosystems, but Pasquale (2015) shows that firms can satisfy transparency formally while defeating intelligibility (e.g., through selective disclosure, complexity, or NDAs). Ananny and Crawford (2018) recommend building institutional conditions for understanding - independence, replicability, and the ability to publish. Governance of visibility identifies the infrastructural locus (interfaces and associated rules) where such institutional guarantees must attach if transparency is to be meaningful.

4. Counter-infrastructures

This paper theorizes counter-infrastructure as any sociotechnical arrangement that seeks to *redistribute observability, control, and the capacity to intervene* away from incumbent platforms and toward publics, researchers, or communities. Building on infrastructure studies, counter-infrastructures can be read as attempts at infrastructural inversion - making the pipes and protocols of platformized communication newly visible and governable (Bowker & Star, 2000; Star, 1999). They also resonate with work on protocol politics, recursive publics in free and open-source cultures (Kelty, 2008), and media-infrastructure scholarship that treats pipes, standards, and gateways as sites of power and contestation (Plantin et al., 2018). In our framework, counter-infrastructures are not merely

alternatives to “interfaces”; they are alternative visibility regimes designed to modify who gets to see platformed publics, with what granularity, latency, and independence.

Two dimensions help organize the landscape. First, dependency on incumbent interfaces: to what extent does a counter-infrastructure rely on the technical and legal affordances of dominant platforms (APIs, ToS carve-outs, tolerated scraping) versus operating on independent rails? Second, governance model: is the arrangement oriented to public interest/commons governance (e.g., scholarly oversight, community norms, protocol standards) or to partisan/ideological enclosure? These axes yield different failure modes vis-à-vis the governance of visibility elaborated earlier: dependency concentrates risk in corporate revocation; independence often trades centralized legibility for fragmented observability; and ideological forks may secure autonomy while narrowing accountability across publics (Benkler et al., 2018).

This paper also proposes three evaluative criteria, aligned with our theoretical aims: (a) independence (freedom from incumbent veto and capture), (b) coverage (ability to observe the arenas where publics actually assemble), and (c) continuity (stability and reproducibility of access over time). No family maximizes all three simultaneously.

Dependent counters operate on incumbent rails—through public APIs, research access programs, or tolerated scraping—to instrument platforms for public oversight. Examples include academic dashboards, watchdog pipelines, and newsroom tools that reconstruct content flows, engagement dynamics, and amplification patterns. Methodologically, they align with audit traditions that infer platform behavior from inputs/outputs (Sandvig et al., 2014), and with computational communication practices adapted to the “post-API” era (Freelon, 2018). Their democratic value lies in comparability and reach: because they observe the dominant arenas of discourse, they can evidence harms and effects at scale. Yet this very dependence exposes a structural fragility. As Bruns (2021) shows, the “APIcalypse” marks a broader turn in which platforms narrowed endpoints, raised costs, or retired data surfaces, shifting researchers from relatively open programmability to contingent, revocable access. Meta’s retirement of CrowdTangle - long a de facto public observability instrument - and Twitter/X’s paywalled API exemplify how visibility can be rescinded unilaterally, reasserting what we elsewhere term *infrastructural capture* (Bruns, 2021; Freelon, 2018).

Within the governance of visibility, dependent counters thereby sit at a paradoxical nexus: they can deliver high coverage and strong continuity *until* a policy or pricing change converts visibility into a corporate asset. Mitigations—legal entitlements to data (e.g., vetted researcher access regimes), redundancy across data sources, and publication independence - seek to re-publicize this visibility, but until such guarantees are enforceable, dependent counters remain foreclosure-prone (Ananny & Crawford, 2018; Pasquale, 2015).

Protocol counters relocate social networking off incumbent allowances and onto open, federated structures, thereby redistributing infrastructural chokepoints. Two prominent designs illustrate the strategy. ActivityPub (a W3C Recommendation) powers the Fediverse (e.g., Mastodon) via client-server and server-server federation, with governance distributed at the instance level through community norms, codes of conduct, and block/allow lists (W3C, 2018). AT Protocol/Bluesky modularizes the stack into Personal Data Servers (identity/graph portability), relays (a “firehose” of events), app views and custom feeds, and labelers for distributed moderation, explicitly designing for forkability and exit (Bluesky, 2024). The theoretical import is twofold. First, protocols instantiate governance-by-design: rules are embedded in standards and reference implementations rather than centralized corporate policy (fPlantin et al., 2018). Second, they relocate visibility governance from a single platform to a

mesh of actors - instances, relay operators, client developers, and standards bodies—potentially reducing the risk of unilateral foreclosure.

However, autonomy introduces observability dilemmas. In the Fediverse, logs are fragmented across instances; federation is selective and dynamic; and research ethics (consent, scraping norms, cross-instance aggregation) vary and remain contested (Wähner et al., 2024). In AT Protocol’s early federation, the practical topology of data flows is shaped by which relays are active and which allowlists are in force, raising questions about where new chokepoints emerge (Bluesky, 2024). Protocol counters thus tend to maximize independence and, once mature, continuity, but trade away canonical telemetry: there is no single vantage point from which to reconstruct system-wide dynamics.

For the governance of visibility, the implication is that observability must be co-governed—through shared research feeds, common event schemas, opt-in aggregation, and federated ethics—rather than dictated by a single firm. In short, protocolization resists enclosure but demands new institutions of measurement and consent (Plantin et al., 2018; Wähner et al., 2024).

A third family comprises ideological or forked counters (e.g., Gab, Truth Social): autonomous stacks that minimize dependency on incumbents by building parallel hosting, moderation, and distribution infrastructures. In infrastructural terms, they succeed at independence and often continuity (no incumbent can “turn off” access). Yet they typically produce sectarian visibility—publics that are internally legible while opaque across ideological boundaries—thereby narrowing accountability and complicating cross-public verification (Benkler et al., 2018; Marwick, 2018). The literature on asymmetric ecosystems shows how such enclaves can stabilize parallel knowledge infrastructures with distinct epistemic authorities, amplification networks, and incentives, reducing the *bridge visibility* needed for democratic contestation (Benkler et al., 2018). From our perspective, these counters demonstrate that autonomy is not equivalent to publicness: infrastructural independence can coexist with visibility narrowing that undermines the broader conditions of accountability.

5. Legislative paths

The DSA Article 40 delegated act aims to convert data access for vetted researchers from benevolence to entitlement, clarifying documentation, formats, and the role of Digital Services Coordinators, and announcing a data-access portal (European Commission, 2025). To align with our framework, three guardrails are crucial: independence (no NDAs on core findings; freedom to publish), reproducibility (versioned datasets; changelogs), and timeliness (latency limits; service-level expectations enforceable by regulators). Implementation risks - vetting criteria, privacy safeguards, platform under-delivery - will determine whether Article 40 yields intelligible transparency rather than compliance theater (Ananny & Crawford, 2018; Pasquale, 2015). For protocols, communities should co-design opt-in research feeds, ethical guidelines for cross-instance aggregation, and mechanisms that respect blocklists while enabling federated observability (Wähner et al., 2024; W3C, 2018).

Public funders should back independent observatories and data trusts not reliant on platform grants. Journals should require methodological disclosures on data provenance, access constraints, and versioning so reviewers can assess dependency risks. Civil society coalitions can align transparency demands across jurisdictions and election cycles, amplifying accountability beyond single-firm arenas (Gorwa, 2019; Pasquale, 2015).

6. Conclusion

Treating social platforms as infrastructures clarifies that governance is not only about content rules or ranking tweaks; it is fundamentally about who is permitted to see the system at scale, at what cadence

and cost, and with what independence to publish. API paywalls, tool deprecations, vetting programs, and policy churn are infrastructural levers by which platforms govern visibility, privatizing the means of observation and, with it, much of the capacity for public oversight.

Counter-infrastructures exist and matter. Dependent dashboards scale oversight but are foreclosure-prone; protocol alternatives resist foreclosure but complicate observability; ideological forks illustrate how infrastructural independence can coincide with narrowed public accountability. Regulation - exemplified by the DSA Article 40 regime - points toward a public reassertion over visibility governance, but realization will hinge on whether intelligibility (independent, reproducible, timely access) is achieved in practice.

Taken together, the three families reveal a trilemma for counter-infrastructures under digital capitalism: one can reliably maximize at most two of independence, coverage, and continuity. Dependent counters deliver coverage and continuity until infrastructural capture revokes access; protocol counters deliver independence and, eventually, continuity but face coverage and legibility challenges absent shared observability primitives; ideological forks deliver independence and continuity within enclaves while sacrificing cross-public accountability. Recognizing this trilemma clarifies why policy and standards—from enforceable research-access rights to protocol-level observability conventions—are necessary complements to technical design. Counter-infrastructures can redistribute who gets to see; only institutional guarantees can stabilize that redistribution over time (Ananny & Crawford, 2018; Pasquale, 2015; Plantin et al., 2018).

The core theoretical move is to center governance of visibility - the infrastructural allocation of the right to know - within digital capitalism. If visibility remains privatized, oversight persists as a patchwork of keyholes. If visibility is re-publicized, through enforceable access rights, technical observability primitives, and independent institutions, we can rebuild the epistemic conditions for democratic accountability in a platformed public sphere.

Funding statement

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



Funded by the
European Union

References

- Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, 20(3), 973–989. <https://doi.org/10.1177/1461444816676645>
- Benkler, Y., Faris, R., & Roberts, H. (2018). *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford University Press.

- Bluesky PBC. (2024). *AT Protocol: Federation architecture* (developer documentation). <https://docs.bsky.app/>
- Bowker, G. C., & Star, S. L. (2000). *Sorting things out: Classification and its consequences*. MIT press.
- Bruns, A. (2021). After the 'APIcalypse': Social media platforms and their fight against critical scholarly research. *Disinformation and data lockdown on social platforms*, 14-36.
- Couldry, N., & Mejias, U. A. (2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford University Press.
- European Commission. (2025, July 2). *Delegated act on data access under the Digital Services Act (DSA)* [Commission Delegated Regulation C(2025) 4340 final]. <https://digital-strategy.ec.europa.eu/en/library/delegated-act-data-access-under-digital-services-act-dsa>
- Freelon, D. (2018). Computational research in the post-API age. *Political Communication*, 35(4), 665–668. <https://doi.org/10.1080/10584609.2018.1477506>
- Gillespie, T. (2010). The politics of “platforms.” *New Media & Society*, 12(3), 347–364. <https://doi.org/10.1177/1461444809342738>
- Gorwa, R. (2019). What is platform governance? *Information, Communication & Society*, 22(6), 854–871. <https://doi.org/10.1080/1369118X.2019.1573914>
- Kelty, C. M. (2008). *Two bits: The cultural significance of free software*. Duke University Press.
- Marwick, A. E. (2018). Why do people share fake news? A sociotechnical model of media effects. *Georgetown Law Technology Review*, 2(2), 474–512.
- Ortutay, B. (2024, August 14). Meta kills off CrowdTangle despite pleas from researchers and journalists. *AP News*.
- Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press.
- Plantin, J.-C., Lagoze, C., Edwards, P. N., & Sandvig, C. (2018). Infrastructure studies meet platform studies in the age of Google and Facebook. *New Media & Society*, 20(1), 293–310. <https://doi.org/10.1177/1461444816661553>
- Sandvig, C., Hamilton, K., Karahalios, K., & Langbort, C. (2014, May). Auditing algorithms: Research methods for detecting discrimination on internet platforms. In *Data and Discrimination: Converting Critical Concerns into Productive Inquiry* (ICA 2014 preconference), Seattle, WA. <http://www-personal.umich.edu/~csandvig/research/Auditing%20Algorithms%20--%20Sandvig%20--%20ICA%202014%20Data%20and%20Discrimination%20Preconference.pdf>
- Srnicek, N. (2017). *Platform capitalism*. Polity.
- Star, S. L. (1999). The ethnography of infrastructure. *American Behavioral Scientist*, 43(3), 377–391. <https://doi.org/10.1177/00027649921955326>
- Starbird, K. (2019). Disinformation's spread: Bots, trolls and all of us. *Nature*, 571(7766), 449. <https://doi.org/10.1038/d41586-019-02235-x>
- Gotfredsen, S. G., & Dowling, K. (2024, July 9). Meta is getting rid of CrowdTangle—and its replacement isn't as transparent or accessible. *Columbia Journalism*. https://www.cjr.org/tow_center/meta-is-getting-rid-of-crowdtangle.php

Tromble, R. (2021). Where have all the data gone? A critical reflection on academic digital research in the post-API age. *Social Media+ Society*, 7(1), 2056305121988929.

W3C. (2018). *ActivityPub* (W3C Recommendation). <https://www.w3.org/TR/activitypub/>

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Profile Books.

Calma, J. (2023, May 31). Twitter just closed the book on academic research. *The Verge*. <https://www.theverge.com/2023/5/31/23739084/twitter-elon-musk-api-policy-chilling-academic-research>

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

DIGIPOL – Deliverable D10

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

Funding statement

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



Funded by the
European Union

Foreword

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

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

Introduction

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

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

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

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

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

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

1. Recognize platform observability as a public-interest infrastructure

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

2. Move from discretionary access to enforceable access pathways

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

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

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

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

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

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

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

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

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

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

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

7. Shift the policy benchmark from transparency to auditability

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

8. Enable explanation systems to support oversight

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

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

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

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

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

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