The Soap Box as a Black Box

Regulating Transparency in Social Media Recommender Systems

Leerssen, P.J.

Publication date
2020

Published in
European Journal of Law and Technology

License
CC BY

The Soap Box as a Black Box: Regulating Transparency in Social Media Recommender Systems

Paddy Leerssen∗

Abstract

Social media recommender systems play a central role in determining what content is seen online, and what remains hidden. As a point of control for media governance, they are subject to intense controversy and, increasingly, regulation by European policymakers. A recurring theme in such efforts is transparency, but this is an ambiguous concept that can be implemented in various ways depending on the types of accountability one envisages. This paper maps and critiques the various efforts at regulating social media recommendation transparency in Europe, and the types of accountability they pursue.

This paper identifies three different categories of disclosure rules in recent policymaking: (1) user-facing disclaimers, (2) government auditing, and (3) data-sharing partnerships with academia and civil society. Despite their limitations and pitfalls, it is argued, each of these approaches has a potential added value for media governance as part of a tiered, varied landscape of transparency rules. However, an important element is missing - public data access. Current trends emphasise exclusive data access regimes directed at particular, trusted regulators or researchers, but this approach has important limitations in terms of scalability, inclusiveness, and independence. This paper articulates the distinct benefits of public data access as a supplement to existing transparency measures, and suggests starting points for its design and regulation.

∗ PhD Candidate, Institute for Information Law (IViR), University of Amsterdam. Non-residential fellow, Stanford Center for Internet and Society. p.j.leerssen@uva.nl.

1   Introduction

Social media platforms have become central actors in media governance. One of their most powerful means of influence is their content recommender systems, which determine the ranking of content as it is presented to users. Their design can therefore have significant effects on what is seen online, and what remains hidden. Accordingly, content recommender systems have a gatekeeping function, implicating urgent public interests and swiftly becoming a key point of control and contention in ongoing debates about online content regulation.[1]

In this otherwise contentious debate, a rare point of consensus for both scholars and policymakers appears to be the need for greater transparency. At present, social media recommendation systems operate largely as 'black boxes", guided by complex, confidential machine-learning algorithms whose operations are inscrutable to outside observers.[2] "A system must be understood to be governed", as Mike Ananny and Kate Crawford observe, and there is broad agreement amongst scholars and policymakers that recommender systems must be more transparent—if not as a sufficient condition for holding them accountable then at least as a first step.[3]

This paper analyses recent policymaking in Europe that attempts to regulate transparency in social media content recommendations. Not yet a cohesive framework, we see various overlapping standards at the national, EU and Council of Europe level, each furthering particular visions of "transparency" and the types of accountability it should serve. This paper critiques these various efforts, drawing on critical literature on transparency regulation and platform governance, and questions how, and under which conditions, they can contribute to holding social media recommender systems accountable for their impact on online information flows.

---

[1] Jennifer Cobbe & Jatinder Singh (2019), 'Regulating Recommending: Motivations, Considerations, and Principles', *European Journal of Law and Technology* 10:3.

[2] Frank Pasquale (2015), *The Black Box Society: The secret algorithms that control money and information*, Harvard University Press.

[3] Mike Ananny & Kate Crawford (2018), 'Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability', *New Media & Society* 20:3.

The paper proceeds as follows: Section 2 offers an overview of governance debates about gatekeeping through social media recommenders, and how these have given rise to calls for greater transparency. Building on recent literature from communications and media studies, it is argued that transparency in social media recommendations is a multifaceted issue which relates not only to the algorithm involved but more broadly to the design and operation of content recommenders as sociotechnical systems. Section 3 describes recent European policymaking around recommendation transparency, identifying three general categories of disclosure rules: user-facing disclaimers, government oversight and civil society partnerships. Each of these methods can have an added value for media governance, despite their respective limitations and pitfalls, as part of a tiered, variegated approach to transparency. Yet, an important element is missing: public disclosures. Section 4 articulates the distinct advantages of public disclosures as a supplement to existing transparency measures, and suggests starting points for their design and regulation.

2 Social media recommenders systems as opaque gatekeepers for online content

2.1 What are social media recommender systems?

Platforms use recommender systems to determine the manner in which content is presented to their users. Their recommendations typically take the form of pages or lists, often referred to as 'feeds', in which the order of content is determined by ranking algorithms. These ranking algorithms can take any number of forms, from simple reverse chronology to complex machine-learning solutions. Recommender systems can also include user customisation options, such as the ability to "like" or "follow" specific content sources, to block or filter certain content sources, or to switch between entirely different ranking logics. Recommender systems are commonly understood as optimising for user attention, or "relevance", but in practice, as will be unpacked further below, recommender design is also shaped by other economic and political imperatives.

Recommendations are not the only way to access social media content. Users can typically also reach content through search functions, user profiles, hotlinking and embedding. Nonetheless, recommender systems can be highly influential, since they commonly take

up a central position in platform interfaces: Facebook's Newsfeed and YouTube's Autoplay and Recommended Videos, for instance, are some of the key content discovery features on their respective platforms. YouTube recently stated that 70% of user viewing is accessed through recommendations.[4] Facebook's Newsfeed being even more central to the platform's interface, the percentage here could plausibly be even higher.

It is important to note that content recommendations are not fully controlled by their operators, but are co-determined by platform users, who influence outcomes in several ways. Firstly, users are responsible for uploading content from which content recommenders draw their recommendations. Secondly, users' behaviour provides feedback signals, including explicit feedback such as rating, following or subscribing, as well as implicit feedback such as scrolling and clicking.[5] Since recommender systems commonly rely on machine-learning processes to optimise the algorithm, these user signals can also serve to shape the weighting of the algorithm over time. Conversely, the recommender system can also shape users' behaviour over time, in terms of their preferences, habits and expectations they form in relation to the service. These complex interactions between the recommendation algorithm and its users make for a recursive and unpredictable system, with the potential for unexpected feedback loops and path dependencies. A notable example is Rebecca Lewis' study of far-right content on Youtube, which emphasises the role of well-organised "influence networks" of content creators and audiences, who used guest appearances and other forms of referral and collaboration to create a pipeline or "rabbit hole" of gradually escalating extremism.[6]

---

[4] Karen Hao (2019), 'YouTube is experimenting with ways to make its algorithm even more addictive', *MIT Technology Review*. Available at: https://www.technologyreview.com/s/614432/youtube-algorithm-gets-more-addictive/

[5] Charu Aggarwal (2016), *Recommender Systems: The textbook*, Springer. Eytan Bakshy, Solomon Messing & Lada Adamic (2015), 'Exposure to ideologically diverse news and opinion on Facebook', 348 *Science* 6239.
David Lumb (2015), 'Why scientists are upset about the Facebook Filter Bubble story', *Fast Company*. Available at: http://www.fastcompany.com/3046111/fast-feed/why-scientists-are-upset-over-the-facebook-filter-bubble-study

[6] Rebecca Lewis (2018), 'Alternative Influence: Broadcasting the Reactionary Right on YouTube', Data & Society Research Institute. Available at: https://datasociety.net/wp-content/uploads/2018/09/DS_Alternative_Influence.pdf

Due to the central role of user behaviour in steering recommendation outcomes, platform recommendations are not fully pre-determined or controlled by their operators. For this reason, communications research into recommender systems has emphasised the importance of looking past algorithms as such towards understanding the complex interactions between technology and its users. Kevin Munger & Joseph Philips warn that decontextualised or monocausal understandings of "the algorithm" shaping online media consumption overestimates the role of their designers and undervalues the relative influence of user communities which shape content supply and demand.[7] Philip Napoli instead characterises gatekeeping on social media as a process of "individual media users working in conjunction with content recommendation algorithms."[8] Building on such insights, Bernhard Rieder, Ariadna Matamoros-Fernandez and Oscar Coromina argue for a shift from studying ranking algorithms to "ranking cultures", acknowledging "the realities of an intricate mesh of mutually constitutive agencies".[9] A more complete understanding of social media recommendations, then, cannot focus on recommendation algorithms alone but must seek to understand the sociotechnical system through which they are produced.[10] Rather, as Natali Helberger, Katharina Kleinen-Vön Königslöw and Rob von der Noll observe, governance of these systems therefore requires close attention to "the complex dynamics between the gatekeepers and the gated".[11] What this sociotechnical perspective demands in terms of transparency will be explored in section 2.3 below, but not before discussing the political economy of recommender system governance that has given rise to such calls for transparency.

---

[7] Kevin Munger & Joseph Philips (2019), 'A supply and demand framework for youtube politics', Working draft. Available at: https://osf.io/73jys/download

[8] Philip Napoli (2015), 'Social Media and the Public Interest: Governance of News Platforms in the Realm of Individual and Algorithmic Gatekeepers', *Telecommunications Policy* 39:1.

[9] Bernhard Rieder, Ariadna Matamoroz-Fernarndez & Oscar Coromina (2018), 'From ranking algorithms to 'ranking cultures': Investigating the modulation of visibility in YouTube search results', *Convergence: The International Journal of Research into New Media Technologies* 24:1.

[10] Rebecca Lewis (2020), 'All of YouTube, Not Just the Algorithm, is a Far-Right Propaganda Machine', *FFWD*. Available at: https://ffwd.medium.com/all-of-youtube-not-just-the-algorithm-is-a-far-right-propaganda-machine-29b07b12430

[11] Natali Helberger, Katharin Kleinen-Von Königslöw & Rob van der Noll (2015), 'Regulating the new information intermediaries as gatekeepers of information diversity', *info* 17:6.

2.2   Recommendation governance: From the attention economy to attention politics

Given their role in shaping online media consumption, content recommendations from dominant social media platforms exercise an important gatekeeping function with implications for online freedom of expression and media pluralism.[12] Their design, which typically optimises for user engagement, stands widely accused of surfacing harmful content and distorting online discourse. In response, social media platforms are now increasingly being pressured by policymakers in Europe and elsewhere to curate their recommendations on the basis of various public interest standards, which has in turn raised concerns about the potential for censorship. The following section describes this emergent political economy of recommender governance in greater detail.

As sites of information gatekeeping, recommender systems invite comparison with editorial decisions in the mass media: they both reflect a subjective (and typically commercially motivated) judgement on what content is "relevant" to their audience.[13] Where they differ is that content recommendations do not determine *access* to content, like a traditional editor would, but rather *exposure*—a function that Helberger, Kleinen-Von Königslöw and Von der Noll describe as "indirect editorial influence".[14] Gillespie makes a similar comparison: "This may be a gentler intervention than an editor deciding what is a front page story and what isn't worth reporting at all, but it is selection nonetheless, and it matters in many of the same ways".[15]

---

[12] On online gatekeeping, see e.g.: Jonathan Zittrain (2006), 'A History of Online Gatekeeping', 19 *Harvard Journal of Law and Technology* 253 (2006). Laidlaw (2015), *Regulating Speech in Cyberspace: Gatekeepers, Human Rights and Corporate Responsibility*, Cambridge University Press. On the specific role of recommendation algorithms in online gatekeeping, see: Helberger, Kleinen-Von Königslöw & Van der Noll (2015).

[13] Tarleton Gillespie (2015), 'Platforms Intervene', *Social Media + Society* 1:1. In the context of search engines, see also: Joris van Hoboken (2012), *Search Engine Freedom: On the Implications of the Right to Freedom of Expression for the Legal Governance of Web Search Engines*. Kluwer International. Van Couvering (2010) Search Engine Bias: The Structuration of Traffic on the World Wide Web. PhD thesis, The London School of Economics and Political Science (LSE).

[14] Helberger, Kleinen-Von Königslöw & Van der Noll (2015), n. 12.

[15] Gillespie (2015), n. 14.

Another distinction with traditional media gatekeeping is that platforms tend to process user-generated content, rather than editorially selected content. Even when media organisations use algorithms to personalise content selections, as the New York Times does for instance, they are still drawing from a smaller pool of vetted content than, for instance, YouTube's Recommended Videos. In this regard, Jennifer Cobbe & Jatinder Singh distinguish "open recommending" of user-generated content by platform services, which is the focus of this paper, from "curated recommending" of walled garden services such as Netflix, or "closed recommending" of in-house content by media organisations such as the New York Times.[16] Whilst all these services use complex algorithmic systems to generate personalised content recommendations, the "open recommending" performed with user-generated content operates at the largest scale and with the greatest diversity of content, serving an essential or even quasi-infrastructural role in many media ecosystems.[17] Given their open nature, they also offer the greatest risk of surfacing harmful or illegal content. In this light, social media recommender systems afford a form of gatekeeping which may at first seem relatively indirect and light-touch, but, given the influential position of a handful of social media platforms, nonetheless has the potential for systemic effects across online media ecosystems. [18]

Social media recommender systems also operate within different organisational and commercial structures than the mass media's editorial selections.[19] Unconstrained by professional and organisational standards of journalism, social media platforms are incentivised to optimise their recommendations primarily for *engagement*.[20] This "attention economy" logic of recommender systems has drawn extensive criticism from

---

[16] Cobbe & Sing (2019), n. 2.

[17] Ben Wagner (2018), 'Free Expression? Dominant information intermediaries as arbiters of internet speech'. In: Martin Moore & Damian Tambini, *Digital Dominance: The Power of Google, Amazon, Facebook, and Apple*, Oxford University Press. José van Dijck, Thomas Poell & Martijn de Waal (2018), *The Platform Society: Public Values in a Connective World*, Oxford University Press. Ethan Zuckerman (2020), 'The Case for Digital Public Infrastructure', Knight First Amendment Institute. https://knightcolumbia.org/content/the-case-for-digital-public-infrastructure

[18] Cobbe & Singh (2019), n.2., describe open recommenders as the 'the biggest contributor to systemic issues' in online speech governance, compared to curated and closed recommenders.

[19] Napoli (2015), n. 9.

[20] Napoli (2015), n. 9. Van Dijck, Poell & De Waal (2018), n. 18.

academia, the press, and policymakers, who have highlighted the potential harms that may arise from recommendations optimised for engagement and are increasingly forcing platforms to incorporate alternative design values.

Engagement-optimised social media recommendations are alleged to contribute to a range of harms (though some critiques have stronger empirical grounding than others). To name only a few: content recommenders have been accused of accelerating extremist content and disinformation[21]; polarising audiences and pushing users into homogenous "filter bubbles" or "echo chambers"[22]; underserving content on certain social movements and news events, or from particular political viewpoints[23]; exposing children and other vulnerable groups to harmful content[24]; and for reflecting or amplifying societal prejudices and biases against marginalised groups.[25] They have also been accused of intentional political bias and censorship, where platforms allegedly intervened with content on specific issues—though many of these claims remain unverified.[26] Such critiques serve to

---

[21] Lewis (2018), n. 7.

[22] Eli Pariser (2011), *The Filter Bubble: What the internet is hiding from you*, Penguin. Cass Sunstein (2001). *Republic.com*, Princeton University Press. For more critical approaches, see: Axel Bruns (2019), *Are Filter Bubbles Real?* Polity Press. Frederik Zuiderveen Borgesius (2016), 'Should we worry about filter bubbles?', *Internet Policy Review* 5:1.

[23] Zeynep Tufekci (2017), *Twitter and Tear Gas: The power and fragility of networked protest*, Yale University Press 2017, p. 156 et seq. (describing how the Ferguson protest movement #Blacklivesmatter in the US 'was almost tripped up by Facebook's algorithm'). For further analysis, also highlighting the ways in which social movements can successfully leverage social media, see: Thomas Poell & Jose van Dijck (2017), 'Social Media and New Protest Movements', in: Jean Burgess, Alice Marwick & Thomas Poell (eds.), *The SAGE Handbook of Social Media*, Sage Press.

[24] James Bridle (2017), 'Something is wrong on the internet', *Medium*. Available at: https://medium.com/@jamesbridle/something-is-wrong-on-the-internet-c39c471271d2. Max Fisher & Amanda Taub (2019) , 'On YouTube's Digital Playground, an Open Gate for Pedophiles', *The New York Times*. Available at: https://www.nytimes.com/2019/06/03/world/americas/youtube-pedophiles.html

[25] Safiya Umoya Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism* (NYU Press 2018).

[26] Thomas Poell & Jose van Dijck (2018), 'Social Media and new protest movements'. In: Jean Burgess, Alice Marwick & Thomas Poell (eds.), *The SAGE Handbook of Social Media* (Sage Press). More broadly, on popular understandings of social media recommendation, see:, Motahhare Eslami et al, (2016). 'First I "like" it, then I hide it: Folk Theories of Social Feeds', *CHI '16*.

problematise and politicise the supposed neutrality or objectivity of platform information flows and their determinations of relevance.[27]

Alternative design principles for social media recommendations are now being devised, and, increasingly, implemented in practice. Academics have articulated a range of different values for content recommenders, including "serendipity"[28], "diversity"[29], "neutrality"[30], "user choice" and "user control"[31], and "agonism".[32] Each reflects different judgements about the particular risks and opportunities posed by recommender systems, and can be operationalised in countless different ways. But what these proposals have in common, is that they depart from the commercial logics of the attention economy, and instead would have social media recommenders reflect public interests or values.[33] Indeed, several governments across Europe have over the past years proposed to regulate social media recommendations through public law, based on a variety of public interest principles and definitions.[34] And platforms are starting to take note.

Since 2016, major social media platforms claim to have altered their recommender systems in ways that depart from a strictly engagement-driven design, ostensibly in response to concerns over the spread of harmful content. In particular, these changes tend to address content which is not explicitly prohibited by the platform but is nonetheless considered undesirable or unwelcome, such as disinformation and political extremism. Facebook in particular has announced a bevy of such measures. In early 2018, the platform changed their Newsfeed recommendation algorithm to promote content shared by friends and

---

[27] For instance: Evgeny Morozov (2014), *To Save Everything, Click Here*, Public Affairs 2014, Ch. 5. ('The perils of algorithmic gatekeeping').
[28] Natali Helberger (2011), 'Diversity by Design', *Journal of Information Policy* 441.
[29] Natali Helberger, Kari Karppinen & Lucia d'Acunto (2016), 'Exposure diversity as a design principle for recommender systems', *Information, Communication & Society* 21:2.
[30] Frank Pasquale (2016), 'Platform Neutrality: Enhancing Freedom of Expression in Spheres of Private Power'. 17 *Theoretical Inquiries in Law* 487.
[31] Jaron Harambam, Natali Helberger & Joris van Hoboken (2018), 'Democratizing algorithmic news recommenders: how to materialize voice in a technologically saturated media ecosystem', 376 *Philosophical Transactions of the Royal Society* 2133.
[32] Kate Crawford (2016), 'Can an Algorithm be Agonistic? Ten Scenes from Life in Calculated Publics', *Science, Technology, & Human Values* 41:1.
[33] Napoli (2015), n. 9. Van Dijck, Poell, & de Waal (2018), n. 18.
[34] See Section 3.2.

reduce the reach of news pages (presented as a move towards more "meaningful engagement").[35] In 2019, they announced the downranking of anti-vaccination content and other "borderline content" which falls short of violating company policies.[36] In the same year, they also announced a new "Click-gap" program to suppress "low-quality content", which is achieved by analysing the relative popularity of a given item on Facebook compared to its overall web traffic.[37] YouTube also claims to be experimenting intensively with methods to improve recommendation quality and to reduce the spread of harmful and misinforming content. A 2019 blog post claimed that "in the last year alone, we've made hundreds of changes to improve the quality of recommendations".[38] Those concerned with combating harmful speech may welcome these interventions, while those concerned with freedom of expression might balk at them. In any case, these examples highlight how recommender systems are increasingly used as a tool for content curation.

A variety of different methods are in play: some interventions target specific speakers or posts, such as Facebook's downranking of false headlines, whereas more fundamental changes to the algorithm have the potential to affect all rankings across the system. Some interventions are decided on a case-by-case basis by human actors, whereas others are automated to a large degree, such as the blacklisting and whitelisting of accounts, keywords or phrases, or analysis of content metadata as in Facebook's aforementioned

---

[35] Adam Mosseri (2018). 'Bringing People Close Together', *Facebook Newsroom*. Available at: https://newsroom.fb.com/news/2018/01/news-feed-fyi-bringing-people-closer-together/
[36] Monica Bickert (2019). 'Combatting Vaccine Misinformation', *Facebook Newsroom*. Available at: https://newsroom.fb.com/news/2019/03/combatting-vaccine-misinformation/;
Facebook (2018). 'How People Help Fight False News', *Facebook Newsroom*. Available at: https://newsroom.fb.com/news/2018/06/inside-feed-how-people-help-fight-false-news/
Mark Zuckerberg (2018). 'A Blueprint for Content Governance and Enforcement', *Facebook Notes*. Available at: https://www.facebook.com/notes/mark-zuckerberg/a-blueprint-for-content-governance-and-enforcement/10156443129621634/ ; Mosseri (2018).
[37] Guy Rosen (2019). 'Remove, Reduce, Inform: New Steps to Manage Problematic Content', *Facebook Newsroom*. Available at: https://newsroom.fb.com/news/2019/04/remove-reduce-inform-new-steps/
[38] YouTube (2019). 'Continuing our work to improve recommendations on YouTube', *YouTube Official Blog*. Available at: https://youtube.googleblog.com/2019/01/continuing-our-work-to-improve.html
(Announcing that YouTube would "begin reducing recommendations of borderline content and content that could misinform users in harmful ways".)

Click-gap program.[39] In any case, as discussed below, these decisions and their effects are largely opaque to outside stakeholders.

These new forms of curation may be motivated by any number of (perceived) demands or pressures, including political pressures and the threat of government regulation.[40] Social media platforms are embedded in complex governance structures and accountability relationships with a range of different stakeholders: not only governments but also proactive users, civil society actors, and commercial partners may motivate them to intervene in content flows. In any case, it is clear that their actions cannot be explained solely through "attention economy" accounts about engagement optimisation. This is not to deny the commercial, profit-seeking nature of social media platforms, but simply to recognise that their economic self-interest may require them to take into account social and political conditions. In other words: recommendation gatekeeping is not simply a matter of attention economy, but also, and increasingly, of attention politics.

This struggle over the future of recommendation gatekeeping does not appear to have definitive answers or solutions. Public interest concepts such as media pluralism – i.e. the appropriate structure or balance of available media in a given polity – cannot, as Kari Karppinen observers, be "solved" objectively or definitively.[41] Indeed, it is worth noting the tensions between current design proposals, such as "diversity" and "trustworthiness" on the one hand and "non-discrimination" and "neutrality" on the other; whereas the former group would require recommenders to seek out and prioritise certain content, the latter could arguably prohibit such differentiation. We need not expect a consensus on such issues to emerge soon: the recognition is growing that there is no such thing as a neutral recommender system, and what remains is a fundamentally political and value-laden

---

[39] On algorithmic blacklisting and whitelisting, see: Jeff Gary & Ashkan Soltani (2019), 'First Things First: Online Advertising Practices and Their Effects on Platform Speech', Knight First Amendment Institute. Available at:
https://knightcolumbia.org/content/first-things-first-online-advertising-practices-and-their-effects-on-platform-speech
[40] E.g. Damian Tambini, Danilo Leonardi and Chris Marsden (2008), 'The privatisation of censorship: self-regulation and freedom of expression'. In: Damian Tambini, Danilo Leonardi and Chirs Marsden, *Codifying cyberspace: communications self-regulation in the age of internet convergence*, Routledge.
[41] Kari Karppinen (2013), *Rethinking Media Pluralism,* Fordham University Press.

question as to what types of content should be prioritised across different segments of the population. It will likely continue to be contested for the foreseeable future, as a new frontier in media governance.[42]

2.3   "Obscured obscuring": The opacity of social media recommendations

A commonly criticised aspect of recommendation governance is that it is deeply opaque. While it is clear *that* platforms increasingly curate their recommendations for various forms of content regulation, *how* they do so is difficult to observe and understand. Recommender systems are perceived as "black boxes", whose internal logics are inscrutable and their outputs unpredictable, creating a barrier to holding these systems accountable.[43] Gillespie memorably warns against "the *obscured obscuring* of contentious material from the public sphere" (emphasis added), which "raises a new challenge to the dynamics of public contestation and free speech".[44] So what makes these systems so opaque? Their lack of transparency is multifaceted, and results from both technical and legal factors.

Taken in its most basic sense, transparency can be said to refer to "the disclosure of certain information that may not previously have been visible or publicly available".[45] In the context of recommender systems, concerns over transparency often refer to the specific algorithms used to produce recommendations. But other aspects of recommender systems are also opaque, such as the outputs (what recommendations are made?) and inputs (user content & metadata, behavioural data, etc.) In addition, transparency can also refer to the human agents and organisational structures involved in designing and operating this system. At present, many influential content recommenders lack transparency on each of

---

[42] Karppinen (2013), n. 42. Natali Helberger (2019). 'On the democratic role of news recommenders'. *Digital Journalism* 7:8.

[43] Pasquale (2015), n. 3.

[44] Tarleton Gillespie (2018), *Custodians of the internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*, Yale University Press.

[45] Oana Brindusa Albu & Mikkel Flyverblom (2016). 'Organizational Transparency: Conceptualizations, Conditions, and Consequences'. *Business & Society* 58:2. See also: Robert Gorwa & Timothy Garton Ash (2020), 'Democratic Transparency in the Platform Society', in: Nate Persily & Joshua Tucker (eds.), *Social Media and Democracy: The State of the Field and Prospects for Reform*, Cambridge University Press.

these issues—from the algorithm as such to its inputs and outputs and the surrounding institutions.

To start with the recommendation algorithms: these are obscure due to their technical complexity as well as intentional corporate secrecy.[46] Given their scale and complexity, these algorithms are often ill-suited to "human scale comprehension", and it is difficult even for experts to develop concrete, causal explanations for specific outcomes.[47] Some platforms now offer individualised "explanation" features, such as Facebook's Why Am I Seeing This? feature, but such efforts have been criticised for failing to meaningfully describe the full complexity of the algorithm's operations.[48] Platforms could in theory publish their algorithms in full and enable outside study, but they have reasons to keep them confidential. First, platforms commonly argue that recommender system design involves commercially valuable trade secrets.[49] Second, confidentiality of the algorithm may in some cases be necessary to prevent users from "gaming" the system and undermining its gatekeeping function.[50] For instance, if platforms were to publish their keyword blacklists, this could help sophisticated spammers to avoid being downranked in this way. Third, documentation of recommender systems algorithms could in some cases jeopardise the privacy of platform users, if this algorithm was developed on the basis of user profile data.

But a clear view of inputs and outputs is also crucial to understanding recommender systems As discussed in Section 2.1, the sociotechnical perspective on recommender systems highlights that the significance of algorithms is very much contextual, as outputs are co-determined by user behaviour. To understand their functioning in practice, a view on their outputs is therefore necessary. Rieder, Matamoroz-Fernandez and Coromina

---

[46] Jenna Burrell (2016), 'How the machine 'thinks': Understanding opacity in machine learning algorithms', *Big Data & Society* 3(1).
[47] Burrell (2016), n 47.
[48] Lilian Edwards & Michael Veale (2017), 'Slave to the algorithm? Why a 'Right to an Explanation' is probably not the remedy you are looking for', *Duke Law & Technology Review* 16:18.
[49] Burrell (2016), n. 47.
[50] Pasquale (2015), n.3 argues that platform's claims in this context should not be taken at face value; only under certain conditions will transparency enable 'gaming', and platforms' claims may in fact be motivated by proprietary concerns. See also, Nick Diakopoulos (2016), Accountability in Algorithmic Decision Making', *Communications of the ACM* 59(2).

conclude that for the opacity of recommender systems, "access to the mythical source code would not solve this problem."[51] Instead, they argue for research methods focused firstly on the *outcomes* of these systems, in terms of what recommendation patterns are generated on particularl issues and for particular publics, and how these change over time.[52]

But the study of recommender system outputs and outcomes is restricted in several ways, first and foremost as a result of their personalisation. Since each user is served a personalised selection of recommendations, it is difficult for any individual observer to make generalisable conclusions about the outputs of the system as a whole.[53] All we know is our own news feeds; as to what others are seeing, we can only guess. Researchers have attempted to counteract this obscuring effect of personalisation through survey techniques, which mobilise a large number of accounts (either bots or human volunteers) to assemble data about the platform's outputs.[54] One notable example out of many is the German *Datenspende* project, which tracked Google search results during a 2017 election for over 4000 participants.[55] However, even the most ambitious and elaborate of these methods can only provide snapshots, and do not come close to a comprehensive or systemic view of platform traffic flows. Worse still, platforms can and have restricted these processes contractually by way of their Terms of Service, and technically by way of blocking

---

[51] Rieder, Matamoroz-Fernandez & Coromina (2017), n. 158.

[52] Rieder, Matamoroz-Fernandez & Coromina (2017), n. 158.

[53] See e.g. Bodo et al. (2018), 'Tackling the Algorithmic Control Crisis: The Technical, Legal, and Ethical Challenges of Research into Algorithmic Agents', *Yale Journal of Law and Technology* 19.

[54] Eduardo Hargreaves et al (2018), 'Biases in the Facebook News Feed: a Case Study on the Italian Elections', *International Symposium on Foundations of Open Source Intelligence and Security Informatics, In conjunction with IEEE/ACM ASONAM*. Available at: https://hal.inria.fr/hal-01907069/document . Christian Sandvig et al (2014), 'Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms', Paper presented to "Data and Discrimination: Converting Critical Concerns into Productive Inquiry," a preconference at the 64th Annual Meeting of the International Communication Association. Available at: http://www-personal.umich.edu/~csandvig/research/Auditing%20Algorithms%20--%20Sandvig%20--%20ICA%202014%20Data%20and%20Discrimination%20Preconference.pdf

[55] Tobias D. Krafft, Michael Gamer & Katharina A. Zweig (2018), Wer sieht was? Personalisierung, Regionalisierung un die Frage nach der Filterblase in Googles Suchmacschine. Final Report for research Project #DatenSpende: Google und die Bundestagswahl 2017. Available at: https://www.blm.de/files/pdf2/bericht-datenspende---wer-sieht-was-auf-google.pdf;

such tools. For instance, Facebook recently blocked a popular data scraping tool by ProPublica, citing violations of its Terms of Service.[56]

Besides independent surveying, one of the most important sources of data regarding recommender systems has been their public APIs, through which outside researchers can download platform data in bulk. But these have come under significant pressure over the past years. Since the Cambridge Analytica scandal, in which academics helped to leak and abuse large sets of user data from Facebook, important APIs have incurred major restrictions in their functionality. This development, which Axel Bruns describes as the "APIcalypse", has been the demise of many widely-used research tools and methodologies, both commercial and academic.[57] Of course, the quality of API access differs between platforms; for instance, YouTube and Twitter offer relatively generous public research APIs, whereas Instagram's was recently shut down entirely.[58] Regardless of what information is currently available, Deen Freelon warns that, since platforms have no binding obligation to maintain these systems in any consistent manner, the situation is fundamentally precarious: "we find ourselves in a situation where heavy investment in teaching and learning platform-specific methods can be rendered useless overnight".[59]

Through code and through contract, then, platforms are able to obstruct independent study of their recommender systems, leaving even the basic outputs unclear.[60] In this sense, most platform content recommenders are even less transparent than the prototypical "black box": not only is it unclear *why* certain decisions are being made, it is simply unclear what decisions are being made in the first place. This is an important contrast with other prominent debates in algorithmic governance, such as, for instance, judicial sentencing algorithms, where the algorithm may be secret but the ultimate

---

[56] Jeremy B. Merill & Ariana Tobin (2019), 'Facebook Moves to Block Ad Transparency Tools — Including Ours. *ProPublica'*. Available at: https://www.propublica.org/article/facebook-blocks-ad-transparency-tools

[57] Axel Bruns (2019), 'After the 'APIcalypse': Social Media Platforms and Their Fight against Critical Scholarly Research', *Information, Communication & Society* 22:11. Available at: https://www.tandfonline.com/doi/abs/10.1080/1369118X.2019.1637447

[58] E.g. Munger & Philips (2019), n. 8.

[59] Deen Freelon (2018), Computational research in the post-API age. *Political Communication* 35(4). https://www.tandfonline.com/doi/abs/10.1080/10584609.2018.1477506?journalCode=upcp20

[60] Julie Cohen (2020), *Between Truth and Power,* Oxford University Press.

decisions are still a matter of public record.[61] It is also a noteworthy contrast with mass media content distribution of press, radio and television, whose outputs are equally a matter of public record and thus make the editorial line of a given outlet readily identifiable for any and all audience members.[62] The personalised gatekeeping of recommender systems, by contrast, is difficult for any outsider to observe at a systemic level.

A final aspect of opacity relates to the organisations surrounding social media recommender systems, which also tend to be poorly documented. As mentioned, speculation abounds regarding the possibility of human interventions in important content recommender systems, such as YouTube's Trending Videos and Facebook's Newsfeed, but there are few conclusive or authoritative sources of information about these platforms' internal operations. In their absence, conjectural "folk theories" and "algorithmic lore" proliferate.[63] Platforms occasionally disclose specific policies, such as Facebook's aforementioned fact-checking partnerships: this program's policies are outlined on the Facebook website, and fact-checkers are required to publish explanations for each fact-checking decision on their respective websites, known as "reference articles". Unfortunately, there is no central repository of these reference articles, leaving them scattered across dozens of websites without any clear standardisation or comprehensive overview. More fundamentally, these partnered fact-checks are but one example out of many possible downranking interventions, which may not be subject to any clear transparency policies at all. How else are platforms and their affiliates intervening? In the most extreme cases, recommendations may not be automated at all, but instead be

---

[61] On transparency & accountability in judicial sentencing algorithm, see e.g.: J. Angwin (2016), Make Algorithms Accountable, *The New York Times*. https://www.nytimes.com/2016/08/01/opinion/make-algorithms-accountable.html. Alyssa M. Carlson (2017), 'The Need for Transparency in the Age of Predictive Sentencing Algorithms', *Iowa Law Review* 103:303.

[62] This comparison has been made in numerous commentaries on news personalization, e.g.: Bodo et al (2018), n. 54; Yochai Benkler, Robert Faris & Hal Roberts (2018). *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. Oxford University Press. It is worth noting that personalization is also a growing trend in (the online editions of) mass media organizations, posing similar problems for the study of these outlets: Nick Diakopoulos & Michael Koliska (2016), 'Algorithmic Transparency in the News Media' 5 *Digital Journalism* 809. Nick Diakopoulos (209) *Automating the News: How Algorithms are Rewriting the Media,* Harvard University Press.

[63] Sophie Bishop (2020), 'Algorithmic Experts: Selling Algorithmic Lore on YouTube', *New Media + Society* 6:1. Eslami et al (2016), n. 27.

curated entirely by human operators. Facebook's by-now notorious Trending Topics was discovered in 2016 to be manually curated by Facebook staff, rather than by an automated algorithmic process, and these revelations quickly prompted accusations of political bias.[64] As such stories illustrate, the opacity of social media recommendations relates not only to their technical specifications, but also the organisational structures in which they are embedded. In the words of Ananny & Crawford, transparency in algorithmic systems should take into account "not just code and data but an *assemblage* of human and non-human actors."[65] Indeed, as platforms are being pushed to take more proactive and substantive responsibility for recommendation outcomes, human involvement in recommender systems is likely to expand in future.[66]

All this means that the quasi-editorial influence exercised by platform recommendations is difficult for outside stakeholders to study, much less evaluate or hold accountable. How platforms design and adapt their algorithmic systems is effectively hidden from public knowledge, as are the actual outputs and outcomes of their choices.

Though platforms have devised a number of self-regulatory transparency measures, these have broadly failed to assuage criticisms. Relevant efforts include user-facing notices (e.g. Facebook's "Why Am I Seeing This" feature discussed above) as well as data sharing projects with civil society. They tend to be met scepticism for several reasons. Firstly, creating meaningful transparency arguably runs counter to platforms' incentives: they have a commercial interest in monetising traffic data and insights, and thus in keeping this information exclusive, as well as a political interest in avoiding negative publicity.[67] Indeed, Facebook's flagship Social Science One Initiative has been marred with delays and controversies; whilst Facebook cites legal concerns over data protection compliance,

---

[64] Michael Nunez (2016), 'Former Facebook Workers: We Routinely Suppressed Conservative News', *Gizmodo.* Available at: https://gizmodo.com/former-facebook-workers-we-routinely-suppressed-conser-1775461006

[65] Ananny & Crawford (2018), n. 4. See also Kaminski (2020, forthcoming), "Understanding Transparency in Algorithmic Accountability", in: Woodrow Barfield (ed.), *Cambridge Handbook of the Law of Algorithms*, Cambridge University Press.

[66] Section 2.

[67] Bruns (2019), n 58.

others blame a lack of incentives and political will.[68] Even the European Data Protection Supervisor recently argued as much: "It would appear therefore that the reluctance to give access to genuine researchers is motivated not so much by data protection concerns as by the absence of business incentive to invest effort in disclosing or being transparent about the volume and nature of data they control."[69] Such considerations may explain the recent attention for government regulation of transparency.

3   State of play: Regulating recommendation transparency in Europe

The law and policy literature displays a strong consensus around the need for greater transparency in social media governance, particularly as regard content recommender systems.[70] Yet it is also widely acknowledged that "transparency" is an ambiguous concept that can be operationalised in numerous ways, particularly as regards such complex technological phenomena as recommender systems. As Robert Gorwa & Timothy Garton Ash argue, "transparency in practice is deeply political, contested, and oftentimes problematic" – or, more bluntly by Amitai Etzioni, "a form of regulation by other means".[71] The following section reviews European plans to regulate transparency in social media recommender systems, and the types of accountability they pursue.

---

[68] European Data Protection Supervisor (2020), A Preliminary Opinion on data protection and scientific research. Available online: https://edps.europa.eu/sites/edp/files/publication/20-01-06_opinion_research_en.pdf

[69] European Data Protection Supervisor (2020), n. 69.

[70] E.g. Cobbe & Singh (2019), 2. Gillespie (2018), n. 45. Van Dijck, Poel & De Waal (2018), n. 18. Pasquale (2015), n. 3. Gorwa & Garton Ash (2020), n. 46. Daphne Keller & Paddy Leerssen (2020), 'Facts and Where to Find Them: Empirical Research on Internet Platforms and Content Moderation', in: Persily N. and Tucker J (eds.), *Social Media and Democracy: The State of the Field and Prospects for Reform* (Cambridge University Press). Ben Wagner, Krisztina Rozgonyi, Marie-Therese Sekwenz, Jennifer Cobbe and Jatinder Singh (2020), 'Regulating transparency?: Facebook, Twitter and the German Network Enforcement Act', *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. Nicolas Suzor (2018), 'Digital Constitutionalism: Using the Rule of Law to Evaluate the Legitimacy of Governance by Platforms', *Social Media + Society*. Laidlaw (2015), *Regulating Speech in Cyberspace: Gatekeepers, Human Rights and Corporate Responsibility*, Cambridge University Press.

[71] Amitai Etzioni (2010), 'Is Transparency the Best Disinfectant?' *The Journal of Political Philosophy* 18:4. Gorwa & Garton Ash (2020).

Transparency measures can be analysed in numerous ways. A substantial literature of different transparency taxonomies has emerged, with early work focusing on government transparency but later turning to address private actors and, more recently, platforms and algorithmic systems in particular.[72] Transparency has accordingly been conceptualised in terms of its subjects, formats, rationales, timing, effectiveness, and so many other factors.[73] A recurring theme in this literature, which will guide the discussion in this paper, is the question *to whom* transparency is offered. This accords with the common understanding of transparency and accountability as relational concepts, which are defined by the stakeholders they serve.[74] Following Ananny & Crawford, interrogating the relationship between a proposed transparency measure and its intended accountability outcome, must start with the question *to whom* accountability will be rendered.[75] A similar relational focus can also be seen in, for instance, the work of David Weil, Mary Graham and Archon Fung on "targeted transparency".[76] By examining the audiences that transparency measures serve, we can begin to chart the more fundamental visions of platform accountability that inform these measures.

In the past years, European policymakers have undertaken several different initiatives to regulate the transparency of social media recommendations. This by now complex and fragmented landscape includes horizontal instruments, such as competition law and data protection law, which are not tailored to social media governance in particular but may still have some spillover benefits for its purposes. More recently we also see the emergence of several sectoral proposals that lay out a specific vision on transparency for social media recommendations in particular. Most of the latter instruments are rooted in media

---

[72] For recent overviews see Gorwa & Garton Ash (2020), n. 71. Ananny & Crawford (2018), n. 4. Kaminski (2020, forthcoming), n. 66.

[73] E.g. David Heald (2006), "Varieties of Transparency", in: Christopher Hood & David Heald, *Transparency: The Key to Better Governance?*, Oxford University Press.

[74] E.g. Mark Bovens (2007), "Analysing and Assessing Accountability: A Conceptual Framework", *European Law Journal* 13:4; Jerry Mashaw (2006), "Accountability and Institutional Design: Some Thoughts on the Grammar of Governance", Yale Law School, Public Law Working Paper No. 116; Richard Mulgan (2000), "Accountability: An Ever-Expanding Concept?", *Public Administration* 78:3.

[75] Ananny & Crawford (2018), n. 4, citing: Glasser (1989), 'Three views on accountability', in: Everette Dennis, Donald Gillmor & Theodore Glasser (eds), *Media Freedom and Accountability*. Praeger.

[76] David Weil, Mary Graham & Archon Fung (2007), *Full Disclosure: The Perils and Promise of Transparency*. Cambridge University Press.

pluralism policy, but they also target other public interest considerations such as the combating of online disinformation. Despite the variety of rules in play, the transparency measures contained in these instruments can be grouped into three general categories, aimed at three different sets of stakeholders: (1) user-facing disclosures, which aim to channel information towards individual users in order to empower them in relation to the content recommender system, (2) government oversight, which appoints a public entity to monitor recommender systems for compliance with publicly-regulated standards, and (3) partnerships with academia and civil society, which enable these stakeholders to research and critique recommender systems. Each of these is discussed below.

	Disclosures for users (end-users and content providers)	Disclosures for public authorities	Disclosures for academia and civil society
Disclosure Form	Disclaimers, notices, "explanations"	Audits, reporting requirements	Data-sharing partnerships, initiatives, observatories, etc.
Associated Accountability Standard(s)	User choice / revealed preference	Public standard setting (e.g. non-discrimination, pluralism, trustworthiness)	Various/undefined.

Table 1: Typology of disclosure rules for social media recommenders in Europe

At the outset, it should be noted that these different types of transparency are by no means mutually exclusive; rather, they reflect the growing consensus that platform governance

requires a multistakeholder approach.[77] Accordingly, most scholars defend a variegated or tiered approach to transparency and accountability in this space, such as Frank Pasquale's "qualified transparency" model and Andrew Tutt's "Spectrum of Disclosure".[78] As the following section shows, European policy is developing such a tiered approach, in which understandable, simplified information is channelled towards individual end-users, and detailed, sensitive information is shared confidentially with experts in government and civil society.[79]

What appears to distinguish social media from other areas of platform governance, is the growing emphasis on transparency for civil society and academia, engaging what Archon Fung describes as "the civic immune system" and Mark Bovens as social and political accountability (as distinct from legal or administrative accountability).[80] Including these actors may seem relatively uncontroversial, relative to direct command-and-control regulation, and indeed it appears that their inclusion is motivated by the politically sensitive nature of media governance. But it is in defining and institutionalising these notionally independent groups that problems are likely to emerge. Maintaining the inclusiveness and independence of such efforts, and ultimately their legitimacy, necessitates that policymakers should also turn their attention towards developing a robust vision for *public* data access for recommender systems, without restrictions on who can access the data involved.

3.1 User-facing disclaimers

Perhaps the most common approach to regulating transparency in recommender systems is to require disclosures for individual users. The aim of transparency in this context is to

---

[77] E.g. Laidlaw (2015), n. 71. Cobbe & Singh (2019), n. 2. Chris Marsden, Trisha Meyer and Ian Brown (2019), 'Platform values and democratic elections: How can the law regulate digital disinformation?', *Computer Law & Security Review*.

[78] Pasquale (2015), n. 3. Andrew Tutt (2017), *An FDA for Algorithms*, 69 Administrative Law review 83. Tal Zarsky (2013), *Transparent Predictions*, University of Illinois Law Review 1503. For overview and analysis, see: Kaminski (2020, forthcoming).

[79] Pasquale (2015), n. 3.

[80] Bovens (2007), n. 75.

inform users about their available options so as to help them realise their own preferences, appealing to such values as individual autonomy, agency and trust.[81] If platforms fail to do so, users can, in theory, respond by exiting the platform and taking their activity elsewhere. Napoli describes this as the "individualist model" of social media governance, in which platforms are required to "provide an enabling environment in which individual responsibility and autonomy can be realised in relation to the production, dissemination, and consumption of news and information".[82] It should be noted that the category of "users" in the context of social media platforms includes not only the consumers of content, but also the providers of content, ranging from amateur vloggers to professional influencers and major media organisations. With that in mind, user-facing transparency can also appeal to principles of competition, fairness and diversity in online media markets.

This user-facing approach to transparency can be seen in several European instruments. The General Data Protection Regulation (GDPR) grants platform users a bundle of individual rights. Article 5 lists "transparency" as one of the Regulation's key principles, and users are granted a bevy of information and notice rights about personal data processing under Article 12-14. More specifically, under Article 22, data subjects may under certain circumstances have the right to opt out of such automated decisions, and also enjoy a bundle of information rights collectively known as the "right to an explanation".[83]

Given that the GDPR focuses on data protection, rather than media governance or platform gatekeeping per se, the information acquired in this way could be of only tangential relevance to the study of platform gatekeeping. However, expansive interpretations may be possible: Max van Drunen, Natali Helberger and Mariella Bastian have studied this right as it applies to news recommender systems, and conclude that these provisions should be interpreted *contextually* as a means to empower data subjects in their capacity as news

---

[81] Max van Drunen, Natali Helberger & Mariella Bastian (2019), 'Know your algorithm: what media organizations need to explain to their users about news personalization', *International Data Privacy Law* 9:4.

[82] Napoli (2015), n. 9.

[83] Andrew Selbst & Julia Powles (2017), 'Meaningful information and the right to explanation', *International Data Privacy Law* 7(4). Margot Kaminski (2019), 'The Right to Explanation, Explained', *Berkeley Technology Law Journal* 34: 1.

consumers.[84] On this basis, they argue that users of recommender systems are entitled to a range of information about e.g. the parties able to influence editorial decisions, the profiles that the algorithms construct about them, and the algorithm's metrics and factors.[85] In such a reading, the GDPR could in theory be a source of insights about platform gatekeeping decisions. It remains to be seen whether such access rights will find much usage with the average end user and in fact serve as a source of empowerment in practice.

Another relevant horizontal instrument is the Regulation on Promoting Fairness and Transparency for Business Users of Online Intermediation Services (Platform-to-Business Regulation).[86] This Regulation affects a different category of users: not *consumers* of social media content, but rather *producers*, who are granted certain notice rights in relation to recommender systems under Article 5. This provision requires platforms to disclose, *inter alia*, "the characteristics of the goods and services offered to consumers through the online intermediation services or the online search engine".[87] For sophisticated content providers who rely on social media, such as newspapers and other media outlets, this could be an additional way to adapt to changes in recommendation algorithms, and potentially to detect unlawful or abusive forms of discrimination.[88]

New proposals particular to media governance are also emerging. The *Medienstaatsvertrag*, proposed in 2018 by the German broadcast authority, requires media intermediaries to disclose the selection criteria that determine the sorting and presentation of content. These include "the central criteria of aggregation, selection and presentation of content and their weight, including information about the function of the

---

[84] Van Drunen, Helberger & Bastian (2019), n. 82.

[85] Van Drunen, Helberger and Bastian (2019), n. 82, also recommend disclosures about e.g. sources (the origin of the information). However, these arguments appear to focus on recommender systems operated by *news organizations*, which have direct knowledge and control of such elements (see Cobbe & Singh 2018, n. 2). Such disclosures do not appear to be feasible in the context of platforms.

[86] Regulation 2019/1150 of the European Parliament and of the Council of 20 June 2019 on promoting fairness and transparency for business users of online intermediation services ('P2B regulation'). *OJ L 186. ('P2B Regulation').*

[87] P2B regulation, Art. 5.

[88] Alessio Cornia et al (2018), Private Sector News, Social Media Distribution, and Algorithm Change. Reuters Institute Digital News Project Report. Available at:
https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2018-
10/Cornia_Private_Sector_News_FINAL.pdf

algorithms used".[89] Addressed towards end-users, they must be made in "understandable language", and in "in easily recognisable, directly accessible and constantly available formats".[90] Comparable recommendations are made in the EU Code of Practice on Disinformation, which is a co-regulatory instrument signed by Facebook, Google and Twitter under the guidance of the European Commission.[91] These companies must "consider empowering users with tools enabling a customised and interactive online experience so as to facilitate content discovery and access to different news sources representing alternative viewpoints, also providing them with easily-accessible tools to report Disinformation"[92] A number of Council of Europe recommendations also emphasise the importance of informing and empowering users. For instance, Recommendation 2018-1 on media pluralism and transparency of media ownership calls on states to encourage platforms to "provide clear information to users on how to find, access and derive maximum benefit from the wide range of content that is available".[93]

Evidently, the notion that individual user rights should "empower" users vis-à-vis social media recommender systems is widespread in European policy circles. But there are also important limitations to these user-centric approaches, both practical and principled. As a practical matter, informing users about complex systems such as content recommenders is difficult, and not straightforwardly achieved through disclaimers or notices. As stated,

---

[89] Staatsvertrag zur Modernisierung der Medienordnung in Deutschland - Entwurf ('Medienstaatsvertrag'), Art 53(d)(translation mine). Available at: https://www.rlp.de/fileadmin/rlp-stk/pdf-Dateien/Medienpolitik/ModStV_MStV_und_JMStV_2019-12-05_MPK.pdf

[90] Medienstaatsvertrag, Art 53(d) (translation mine).

[91] European Commission (2018), Code of Practice on Disinformation. Available at: https://ec.europa.eu/digital-single-market/en/news/code-practice-disinformation ('EU Code of Practice on Disinformation')
Although this document has been characterized as self-regulatory, it is better understood as co-regulatory given the European Commission's close involvement in its development and implementation. See: Aleksandra Kuczerawy (2020, forthcoming), 'Fighting Online Disinformation: Did the EU Code of Practice Forget about Freedom of Expression?', in: Elzbieta Kużelewska, Georgios Terzis, Daniel Trottier and Dariusz Kloza, *Disinformation and Digital Media as a Challenge for Democracy*, Intersentia.

[92] EU Code of Practice on Disinformation, p. 3.

[93] Council of Europe (2018), Recommendation CM/Rec(2018)1 of the Committee of Ministers to member States on media pluralism and transparency of media ownership: Guidelines on media pluralism and transparency of media ownership. For a critical discussion of this user-centric strand in Council of Europe standards, see: Napoli 2014.

the complexity of recommender systems renders "algorithmic explanations" difficult if not impossible, certainly in formats that are digestible to the average end-user. Evidence from privacy and consumer protection law scholarship shows that user-facing notices on social media platforms and other websites are routinely neglected by the vast majority of users.[94] And even where information is made to be "simplified" and "understandable", as media governance instruments are now requiring, these effects are likely to persist—the most infamous precedent being the cookie consent notices required under EU privacy law.[95]

Even if fully informed, individual users may simply lack the market *power* to depart from dominant platform offerings. Due to such well-documented dynamics as market concentration, network effects, and user-lock in, it may be costly or even impossible for users to switch to viable alternative platforms.[96] In this sense, transparency towards users may not have full effect if it is "disconnected from power" to actually change outcomes.[97]

Given these manifold constraints on user-facing disclosures, it remains debatable whether expanding individual transparency rights will have much impact on the average user. A greater impact might be expected with more sophisticated platform users, such as professional content providers or media organisations who rely on social media to ply their trade. Also worth noting is that academics and journalists are starting to experiment with access rights under the GDPR (exercised directly by the researcher or indirectly with the help of volunteers) as a source of data; as Jef Ausloos argues, individualised user rights may thus have unexpected spillovers from their stated goal of individual empowerment to a more collective and social form of accountability.[98]

---

94 Frederik Zuiderveen Borgesius (2015), 'Behavioural sciences and the regulation of privacy on the internet. In Anne-Lise Sibony & Alberto Alemanno (eds.), *Nudging and the law: what can EU law learn from behavioural sciences?* Hart Publishing.
95 Zuiderveen Borgesius (2015), n. 95.
96 Patrick Barwise and Leo Watkins (2018), *The evolution of digital dominance: how and why we got to GAFA.* In: Martin Moore and Damian Tambini (eds.), *Digital Dominance: The Power of Google, Amazon, Facebook, and Apple*, Oxford University Press.
97 Ananny & Crawford (2018), n. 4.
98 Jef Ausloos (2019), 'GDPR Transparency as a Research Method'. SSRN Working paper. Available online at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3465680

More fundamentally, the ideal of "user empowerment" can be criticised as overly individualistic, and endorsing a "neoliberal model of agency".[99] While informing users may serve to enhance choice and competition, as Napoli points, media governance has typically not allowed the public interest to be defined exclusively by these market-based ordering principles. The view that "the public's interest, then, defines the public interest"[100] is marginal, certainly in the European tradition. Rather, media policy has also relied on public and collective forms of governance, including government oversight and professional self-regulation, in order to safeguard public values that risk being underserved in a laissez-faire environment, such as pluralism, diversity, child protection, and localism.[101] Of course, individualist values such as choice, autonomy, competition and agency may still be recognised as important within a broader conception of the public interest. But to *equate* them with the public interest, is to oversimplify the challenges of media governance.

3.2   Government oversight

Several European institutions have proposed government oversight of social media recommendations, in order to safeguards public interest principles such as diversity or child protection,[102] enforced by independent regulatory agencies. In terms of transparency, this form of governance relies on reporting duties for platforms and/or auditing powers vested in the regulator.[103] With relevant expertise and the ability to ensure confidentiality of information disclosed, governments can process more detailed information than user-facing notices allow. Government oversight frameworks for social media recommenders are not yet as commonplace in Europe as user-facing disclaimers, but a number of horizontal instruments apply, and several sectoral proposals have surfaced in recent years.

---

[99] Ananny & Crawford (2018), n. 4. Elettra Bietti (2020), 'Consent as a Free Pass: Platform Power and the Limits of the Informational Turn', 40 *Pace Law Review* 307.

[100] Mark S. Fowler & Daniel L. Brenner (1982), 'A Market Place Approach to Broadcast Regulation', 60 Texas Law Review 207.

[101] Napoli (2015), n. 9.

[102] Napoli (2015), n. 9.

[103] Caveat: as Kaminski (2020, forthcoming) observes, governments commonly rely on knowledge sourced from other societal actors. Accordingly, transparency measures aimed at other stakeholders may also offer indirect benefits to government oversight.

One of the most advanced proposals for public oversight of social media recommendations is Germany's aforementioned *Medienstaatsvertrag*. Its key requirement would be non-discrimination. Under this framework, social media platforms "may not unfairly disadvantage (directly or indirectly) or treat differently providers of journalistic editorial content to the extent that the intermediary has potentially a significant influence on their visibility".[104] German broadcast regulators at the federal and local level would be empowered to set detailed standards for social media recommender design, and to request documentation from platforms about their activities.[105] In the Netherlands, the Dutch State Commission on the Parliamentary System has proposed a comparable "independent entity" to monitor social media recommenders, but in contrast to their German counterparts their mandate would focus not on non-discrimination but rather on maintaining "diversity" and avoiding "bias".[106] "If a strong bias can be observed which does not correspond to the information offered by the users themselves on the platform, or if that bias suddenly changes during an election period, this entity can remark on this and ask the company for a response."[107]

At EU level, the main instrument for media regulation is the Audiovisual Media Services Directive. However, it does not contain any particular rules related to recommender systems. More relevant for our purposes is the EU Code of Practice on Disinformation, which requires signatories to "[d]ilute the visibility of disinformation by improving the findability of trustworthy content" an to "invest in technological means to prioritise relevant, authentic, and authoritative information where appropriate in search, feeds, or other automatically ranked distribution channels".[108] However, this Code is a non-binding co-regulatory instrument, and it lacks any concrete sanctions or enforcement mechanisms; platforms were merely expected to self-report their compliance efforts in the months prior

---

[104] Medienstaatsvertrag, Art. 53(e) (translation mine)

[105] Medienstaatsvertrag , Article. 53(f) (translation mine)

[106] Johan Remkes et al. (2018), *Lage Drempels, hoge dijken: Eindrapport Staatscommissie Parlementair Stelsel*, Boon Publishers. Available online at:
https://www.staatscommissieparlementairstelsel.nl/documenten/rapporten/samenvattingen/12/13/eindrapport

[107] Remkes et al. (2018), n. 107.

[108] EU Code of Practice on Disinformation, p. 3.

to the European Election of May 2019. In terms of transparency, then, it is not armed with the same investigative powers as a conventional regulatory agency. Binding regulation at EU level does appear to be under consideration: leaked policy briefs from the Von der Leyen Commission from 2019 envisages "a dedicated regulatory structure" for the oversight of online platforms, with a particular focus on creating transparency.[109]

The Council of Europe has also developed standards on the need for government oversight of content recommenders, emphasising diversity or pluralism as a guiding principle. Their Committee of Ministers has recommended that "[s]tates should encourage social media, media, search and recommendation engines and other intermediaries which use algorithms … to engage in open, independent, transparent and participatory initiatives that seek to improve these distribution processes in order to enhance users' effective exposure to the broadest possible diversity of media content."[110] In contrast to the foregoing examples, this wording does not expressly refer to regulatory agencies but instead describes in more general terms a need for "open" and "participatory" institutions or initiatives, suggesting a more co-regulatory or multistakeholder approach. The state's more modest role lies in "encouraging" such efforts.

Government oversight of platform recommender systems can also be found in horizontal instruments in data protection and competition law. The General Data Protection Regulation sets limits and conditions on the processing of personal data by content recommender systems, which constrains their ability to personalise content. These rules can be enforced privately by data subjects, but also by national data protection authorities (DPAs). Likewise, competition law constrains dominant platforms in their ability to discriminate between commercial actors on their platform, as a potential abuse of their dominant position.[111] This standard is most relevant for vertically integrated platforms,

---

[109] Alexander Fanta & Thomas Rudl (2019). 'Leaked document: EU Commission mulls new law to regulate online platforms'. *Netzpolitik.* Available at: https://netzpolitik.org/2019/leaked-document-eu-commission-mulls-new-law-to-regulate-online-platforms/#spendenleiste

[110] Council of Europe (2018), Recommendation CM/Rec(2018)1 of the Committee of Ministers to member States on media pluralism and transparency of media ownership: Guidelines on media pluralism and transparency of media ownership, Article 2.5.

[111] Treaty on the Functioning of the European Union, Art. 102.

who also produce their own content and thus have an incentive to discriminate against rival content providers.[112] Both data protection and competition authorities are vested with a bevy of investigative powers, such as requesting documentation and performing audits. These frameworks do not directly address the same public interest concerns as media policy, so it is unlikely that these efforts will be targeted directly at studying media governance issues such as pluralism or disinformation. Nonetheless, their research may still have spillover effects between regulatory fields, potentially revealing information that is relevant to media governance.[113]

Government oversight of social media recommendations faces many significant challenges, both practical and principled. Most straightforward is the fact that government authorities are capacity-constrained, particularly as regards the technical expertise required to perform complex algorithmic auditing, and in relation to the sheer scale and scope of potential research issues at stake in social media governance. This is especially true for horizontal agencies such as competition and data protection authorities, for whom social media recommender systems risk being overshadowed and overlooked in an extensive, economy-wide portfolio. Sectoral proposals, on the other hand, would in many cases require the creation of *entirely new* oversight bodies, or for traditional broadcast regulators to develop radically new forms of expertise. What makes this particularly challenging is that, in Europe, media policy is largely a national affair, without a clear institution at EU level capable of performing a monitoring role. Indeed, EU governments have repeatedly shot down proposals for creating a supranational media authority.[114] National-level action in this space, on the other hand, could result in a duplication and fragmentation of efforts.

---

[112] Helberger, Kleinen-Von Konigslow, von der Noll (2015), n. 12.
[113] European Data Protection Supervisor (2016), EDPS Opinion 8/2016 on coherent enforcement of fundamental rights in the age of big data. Available at
https://edps.europa.eu/sites/edp/files/publication/16-09-23_bigdata_opinion_en.pdf
[114] Beata Klimkiewicz (2009), 'Is the Clash of Rationalities Leading Nowhere? Media Pluralism in European Regulatory Policies', in A. Czepek, M. Hellwig and E. Nowak (eds.) *Press Freedom and Pluralism in Europe: Concepts and Conditions*, Intellect.

It is worth nothing that, given these capacity constraints on government monitoring, government agencies commonly rely on knowledge sourced from other societal actors, through such formats as public consultations, expert hearings, and complaint procedures. Therefore, as Margot Kaminski observes, transparency measures aimed at third parties such as users, civil society and other stakeholders can also serve indirectly to enhance accountability to public regulation.[115]

Principled objections to government monitoring as a form of transparency are also possible. As discussed, public standard setting for recommender systems necessarily involves (quasi-)editorial judgements, which are not readily quantifiable or "solvable" in any objective manner.[116] Such editorial judgements in the mass media have historically been protected against direct government regulation, given the threats to freedom of expression,[117] and attempts to regulate recommendations may raise similar concerns. From this perspective, government attempts to prescribe what is downranked risk becoming a form of censorship – and what is promoted, a form of propaganda.[118] How can a government agency make such essentially political assessments in a legitimate and trustworthy manner?

Put differently, government auditing powers continue to raise issues related to what Kaminski terms "second-order accountability": is the governance system itself sufficiently open to outside scrutiny?[119] If government determinations rely on privileged access to confidential data, which is not accessible to broader publics, it may be difficult for citizens to scrutinise and contest government policy in this space.

This critique of second-order accountability is in line with constitutional principles on the rule of law, due process and open government, which reflect broad agreement that government action should be documented publicly inasmuch as possible.[120] Also relevant is the Council of Europe's emphasis that oversight of social media recommendations should

---

[115] Kaminski (2020, forthcoming), n. 66.
[116] See Section 2.2 & 2.3.
[117] E.g. European Court of Human Rights, *Jersild v. Denmark,* Application no. 15890/89.
[118] See e.g. Van Hoboken 2012 in the context of search engines.
[119] Kaminski (2020, forthcoming), n. 66.
[120] On platform governance and rule of law principles, see Suzor (2018), n. 71.

itself be conducted through "open" and "transparent" initiatives.[121] From this perspective, the legitimacy of government action regarding content recommendations depends on its ability to publicise their actions in a meaningful way. However, publicly documenting algorithmic gatekeeping involves significant technical and operational challenges (as discussed in Section 4 below), and has unfortunately not received detailed attention in relevant standards to date.

A final note on the transparency of government relates to *informal* government action. It is by now well-documented in platform governance that governments can and have used informal means of persuasion and coercion, including the *threat* of regulation, to persuade platforms to adopt certain policies—a stratagem also known as "jawboning", "power laundering" or "regulation by raised eyebrow".[122] As a result, it can be difficult to disentangle public and private sources of influence in online content moderation; what is presented as a private platform policy may in fact be inspired or compelled by governments, whose role becomes obscured. Indeed, this informal approach is exemplified in the European Commission's ongoing reliance on quasi-voluntary "Codes".[123] These informal dimensions of public power risk sidestepping safeguards applicable to formal government action, including transparency principles.[124] In this light, transparency obligations focusing solely on formal government action may fail to capture the full picture. This is where independent disclosure obligations imposed on the platforms may be useful: they may offer a useful starting point not only for holding the platform itself accountable, but also for detecting and contesting informal government action.

---

[121] Council of Europe (2018), n. 94.

[122] Respectively: Derek Bambauer (2015), 'Against Jawboning', 100 *Minnesota Law Review* . Daphne Keller (2019), 'Who Do You Sue? State and Platform Hybrid Power over Speech'. Hoover Institution Aegis Series, No 1902. Available at: https://www.hoover.org/sites/default/files/research/docs/who-do-you-sue-state-and-platform-hybrid-power-over-online-speech_0.pdf. Yochai Benkler (2011), 'A Free, Irresponsible Press: Wikileaks and the Battle over the Soul of the Networked Fourth Estate', 46 *Harvard Civil Rights-Civil Liberties Review*. For application to the ECHR context, see: Paddy Leerssen, 'Cut out by the middle man: the free speech implications of social media blocking and banning in the EU'. *JIPITEC* 6:2.

[123] Kuczerawy (2019), n. 92.

[124] Tambini, Leonardi & Marsden (2008), n. 41.

3.3   Research partnerships with academia and civil society

Recent European standards increasingly emphasise the role of independent researchers from academia, civil society, and related categories such as "the research community" or "media organisations". The types of accountability envisaged with these measures are various: in some cases, these actors are formally incorporated in (co-)regulatory decision making processes, and serve clearly designated accountability functions such as fact-checking or regulatory guidance. In other cases, the aims of involving independent researchers appear to be more open-ended, treating independent research and reporting as an end in itself.

A formalised role for civil society actors can be found in in the Council of Europe's 2018 Recommendation on Media Pluralism, which proposes "open, independent, transparent and participatory initiatives by social media, media actors, civil society, academia and other relevant stakeholders" which would be tasked not only with enabling independent research but also with devising new strategies to ensure diversity and other public interest principles in online content distribution.[125] In France, a 2019 report for the Secretary for Digital Affairs similarly recommends a permanent convening of a "political dialogue with social networks involving the regulator and civil society", including "NGOs, regions and the educational and academic communities" with the government tasked with ensuring transparency for the stakeholders involved.[126] Academia and civil society are also increasingly represented in the voluntary self-regulatory organs, ranging from the long-standing Global Network Initiative to Facebook's novel and widely-publicised Oversight Board.[127]

---

[125] Council of Europe (2018), n. 94.

[126] French Secretary of State for Digital Affairs (2019), 'Creating a French framework to make social media platforms more accountable: Acting in France with a European vision', Final Mission Report on the Regulation of social networks. Available at: https://www.numerique.gouv.fr/uploads/Regulation-of-social-networks_Mission-report_ENG.pdf

[127] Robert Gorwa (2019), 'The platform governance triangle: conceptualising the informal regulation of online content', *Internet Policy Review* 8:2.

More open-ended calls to enable independent research can be found in the EU Code of Practice on Disinformation. Its signatories have committed to "empower the research community", which includes "sharing privacy protected datasets, undertaking joint research, or otherwise partnering with academics and civil society organisations if relevant and possible"; and to "convene an annual event to foster discussions within academia, the fact-checking community and members of the value chain".[128] In late 2019, the European Commission also issued a call for tenders for a new European Digital Media Observatory, which would allow "fact-checkers and academic researchers, to bring together their efforts and actively collaborate with media organisations and media literacy experts", with the aim to "fight disinformation online".[129] To this end, the Observatory would also "help design a framework to ensure secure access to platforms' data for academic researchers working to better understand disinformation".[130] The UK's DCMS White Paper would task the government with "encouraging" the creation of "access for independent researchers", with the aim of "ensure that academics have access to company data to undertake research, subject to suitable safeguards" in order to "help the regulator to assess the changing nature of harms and the risks associated with them."[131]

Whether it is for independent research or as part of some more formalised co-regulatory process, all of the transparency arrangements in this space tend to emphasise the sharing of data with specific, selected institutions—as "partners", "initiatives", or "observatories". No explicit attention is paid to creating robust systems of *public* access, available to academia and civil society at large. This preference for partnerships appears to be motivated by the risk of abuse of sensitive data, as highlighted in Cambridge Analytica scandal.[132] By selecting and vetting trustworthy civil society "partners", and imposing

---

[128] EU Code of Practice on Disinformation, p. 1.
[129] European Commission (2019), 'Commission launches call to create the European Digital Media Observatory'. Available at: https://ec.europa.eu/digital-single-market/en/news/commission-launches-call-create-european-digital-media-observatory
[130] European Commission (2019), n. 130.
[131] Department for Digital, Culture, Media & Sport (2019), *Online Harms White Paper*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/7
93360/Online_Harms_White_Paper.pdf
[132] Freelon (2018), n. 60.

binding conditions and potential sanctions on their access to research data, partnerships have a clear utility in enabling research into sensitive data while reducing the risk of its abuse.

But this selecting and vetting of eligible civil society participants brings challenges of its own. Compared to public datasets, one necessarily reduces the number of stakeholders who can access relevant data and perform research, thereby limiting the potential scale and impact of disclosures. More fundamentally, the selection of eligible participants raises difficult questions about the inclusiveness, diversity and independence of the access framework. The Council of Europe recommends that data access initiatives should be "open, independent, and participatory", as mentioned previously, but what does this ideal look like in practice? For academia but especially for civil society in a broader sense, there is a very clear tension between these ideals of openness and inclusiveness on the one hand, and the push to restrict access to trusted participants on the other hand. As will be argued below, European governments will face significant challenges in instituting such social media watchdogs – and public access can help to address relevant concerns.

Academics make promising candidates for research access, not only given their professional expertise and ethical standards, but also due to the university system allowing for a relatively stable and objective means of accreditation (as well as, at the EU level, the European Research Council). Where self-regulatory efforts in this space such as Social Science One been criticised for slow rollout, a lack of (perceived) independence, and a lack of diversity in its leadership, binding regulation could play an important role in addressing relevant concerns and facilitating access for even the most critical research perspectives.[133] This would require neutral and impartial processes for vetting researchers and holding them accountable to data protection and research ethics standards, perhaps in co-regulation with academic institutions themselves. A useful starting point for such efforts is the European Data Protection Supervisor's recent Preliminary Opinion on Data Protection and Scientific Research, which envisages the creation of a co-regulatory accreditation

---

[133] Bruns (2019), 57. De Vreese et al (2019), Public statement from the Co-Chairs and European Advisory Committee of Social Science One. *Social Science One*. Available at: https://socialscience.one/blog/public-statement-european-advisory-committee-social-science-one

scheme and Code of Conduct for research integrity under the guidance of Data Protection Authorities.[134]

However, despite the clear benefits of such academic research frameworks, they cannot make up for the full breadth of civil society watchdog functions in media governance, which have also been performed by journalists, activists, NGOs and political campaigners. For instance, academia tends to have slow turnover times, and may therefore be ill-equipped to perform real-time, large-scale tasks such as, for instance, fact-checking or election monitoring.[135] Activists, journalists and other civil society actors outside the university system are developing powerful new practices such as algorithmic journalism, platform journalism, and social media activism, which risk being excluded in an academics-only approach to social accountability.[136] Yet for these non-academic institutions, whose membership is often porous and fragmented, defining and accrediting eligible participants is even more fraught. Attempts could be made to devise clear and objective processes for accreditation, which, as in academia, could interface with existing self-regulatory bodies in the field of journalism such as the European Federation of Journalists. But even this approach may be at once too broad and too narrow: on the one hand, if the goal is indeed to limit disclosures of confidential data to a restricted group of trusted and accountable actors, such broad professional structures might be overly broad and enable abuse. And on the other hand, these professional structures could still be considered too restrictive since they still exclude a range of non-traditional watchdogs such has citizen journalists, activists, influencers, bloggers or NGOs. In essence, what is at stake here is a tension between the practical need to restrict sensitive data access to vetted actors, and conceptions of the fourth estate and civil society as open and participatory institutions.

---

[134] European Data Protection Supervisor (2020), n. 69.
[135] Rory Cellan-Jones (2017), Facebook's News Feed experiment panics publishers. *BBC News*. Available at: https://www.bbc.com/news/technology-41733119. More generally, see: Cornia et al (2018).
[136] On algorithmic journalism as an emerging field, see: Nick Diakopoulos (2014), Algorithmic Accountability: Journalistic Investigation of Computational Power Structures. *Digital Journalism* 3:3. Nick Diakopoulos, The Algorithms Beat: Angles and Methods for Investigation. In: Jonathan Gray & Liliana Bounegru (eds.), *The Data Journalism Handbook 2.0.*

The difficulties in defining civil society membership are evident in Facebook's self-regulatory attempts to partner with civil society. For instance, Facebook's fact-checking program, which partnered with independent journalists through the Poynter Institutes' International Fact-Checking Network, drew extensive criticism for including the US based Daily Caller as a partner.[137] This website has been accused by many left-leaning outlets of playing a key role in spreading disinformation and hate speech, arguably invalidating their position as a reliable fact-checker.[138] The point was even raised during CEO Mark Zuckerberg's testimony before Congress.[139] Ultimately, the partnership was terminated in November 2019.[140] A comparable controversy occurred when the Weekly Standard, another right-leaning fact-checking partner, rated an article from the left-leaning ThinkProgress as false.[141] Facebook's widely-publicised Oversight Board is no exception to this trend; the announcement of its membership was met with immediate backlash, mainly consisting of conservative press alleging left-wing bias, but also, for instance, broader concerns about a lack of geographical diversity (e.g. an excess of US members; insufficient Southeast Asian members).[142] Such cases illustrate the difficulty, amidst the ongoing decline in trust in mainstream media and established knowledge institutions, of arriving at generally accepted definitions and configurations of "civil society".

---

[137] Aaron Rupar (2019), 'Facebook's controversial fact-checking partnership with a Daily Caller-funded website, explained', *Vox*. Available at: https://www.vox.com/2019/5/2/18522758/facebook-fact-checking-partnership-daily-caller

[138] Robert Faris et al (2017), 'Partisanship, Propaganda, and Disinformation: Online Media and the 2016 U.S. Presidential Election', Berkman Klein Center Research Publication 2017-6. Available at: https://dash.harvard.edu/bitstream/handle/1/33759251/2017-08_electionReport_0.pdf?sequence=9&isAllowed=y

[139] Ali Breland (2019). 'AOC Asked Mark Zuckerberg About Facebook's Fact-Checking Process. He Didn't Give Her the Whole Truth'. *Mother Jones*. Available at: https://www.motherjones.com/politics/2019/10/aoc-zuckerberg-facebook-congress-daily-caller-fact-check-dodge/

[140] Rupar (2019), n. 138.

[141] Zack Beauchamp (2018), 'Facebook blocked the spread of a liberal article because a conservative told it to', *Vox,* Available at: https://www.vox.com/policy-and-politics/2018/9/12/17848026/facebook-thinkprogress-weekly-standard

[142] E.g. Robin Pagnamenta (2020), 'Facebook will rue its left-wing oversight board appointments', *The Telegraph.* Available at: https://www.telegraph.co.uk/technology/2020/05/06/facebook-will-rue-left-wing-oversight-board-appointments/. Jenny Domino (2020), "Why Facebook's Oversight Board is Not Diverse Enough', *Just Security*. Available at: https://www.justsecurity.org/70301/why-facebooks-oversight-board-is-not-diverse-enough/

It is worth noting that the regulation of civil society research access is in a less advanced stage compared to user-facing disclaimers and government oversight; it is largely limited to soft law, and no binding legislation has yet been proposed in this space. It remains to be seen whether and how relevant legislators will take up their cause in upcoming rounds of legislation. Existing standards do indicate, however, a focus on vetted partnerships for privileged data access, as opposed to the furnishing of publicly accessible information.

In sum, the push to create exclusive research access programs has important advantages in enabling in-depth investigative work, but also has limitations. There are important trade-offs between the vetting of eligible researchers for sensitive data access, on the one hand, and the potential scale, diversity and independence of such programs on the other hand. Forthcoming plans for regulated research access will require careful attention to institutional design so as to manage such trade-offs, and ensure the independence and credibility of these research efforts. Overall, these confidential access programs will have much to offer for in-depth academic research, but appear less suitable for real-time monitoring and reporting by journalists, activists and other non-academic civil society actors.

4   The case for public access

The above has shown that the emergent European framework for transparency in social media recommendations focuses on channelling information towards user-facing notices, government authorities, and civil society research partners. In this landscape, preciously little information is made publicly available. Independent observation of personalised recommendations is obstructed by their technical and legal design. User-facing disclosures, whilst public, are typically simplified and individualised. Detailed data is accessible only to a privileged few in government and selective research partnerships.

A robust regime for public access, I argue, would contribute not only to the first-order goal of making oversight of platforms more effective, but also to the second-order goal of

making the governance system as a whole more open to outside critique.[143] This section articulates these potential benefits associated with public access, and suggests some starting points for its design and regulation. In particular, these recommendations focus on the automated, real-time disclosure of high-level, anonymised information about recommendation system outputs, audiences, and organisations.

4.1   The pros and cons of public access

The main drawback of public records, compared to confidential disclosures such as data sharing partnerships and government auditing, is their limitations in sharing sensitive data: public access requires a trustless design that pre-empts abuse by malicious actors. In the context of platform recommender systems, disclosures would need to contend with threats to user privacy, and, according to platforms, the integrity of the service (i.e. by enabling third parties to "game" the algorithm).[144] Privacy-by-design techniques such as anonymisation and differential privacy can go some way in mitigating these concerns. Nonetheless, publicity places hard limits on what can be disclosed and thus on the ultimate research utility of public disclosures.

However, public access also has important advantages over data partnerships in terms of increasing inclusiveness and scalability. By simply making information publicly accessible, one side-steps the pitfalls of needing to define, accredit or otherwise institutionalise such factious and amorphous categories as "civil society" or "academia". Public disclosures would be available to every researcher with the time and interest—not the lucky few with the wherewithal and *bona fides* to engage in protracted negotiations, tender procedures or other forms of partnership arrangements. In particular, public access opens the doors to civil society actors that do not have an institutional means of accreditation, such as many independent journalists, NGOs, activists, and so forth. In this way, public records offer the prospect of broader and more diverse uptake. Public access can also mitigate threats to

---

[143] This distinction between first- and second-order accountability is drawn from the work of Kaminski (2020, forthcoming), n. 66. See also section 3.2.

[144] Burrell (2016), n. 47. Pasquale (2015), n. 3.

researchers' independence, both real and perceived, since it leaves its users free to pursue critical lines of research without needing to appease the purveyors of data—be they platforms in a self-regulatory setting, or governments in a regulated setting. In this way, public records avoid many of the aforementioned problems with more institutionalised "partnership" models for data access.

The above suggests that public records could be instrumental for real-time, high-level monitoring by media watchdogs such as academics, journalists, activists, and NGOs. As discussed in Section 3.3, academics and others performing more in-depth research may be relatively well-served by privileged research partnerships. But even here it is worth considering that public records can offer a low-cost starting point for more in-depth research. For instance, public records may not suffice to conclusively demonstrate bias or discrimination in a recommendation algorithm, but at a minimum they can offer a starting point for such inquiries by rendering visible trends and disparate outcomes in the system's recommendation outputs. Such evidence can then be used to request further clarification from the platform,[145] or investigate with more fine-grained tools, such as algorithmic auditing approaches,[146] data surveys[147] or GPDR data access requests.[148] In other words, public access can serve as a first-warning system for more targeted efforts.

The open nature of public disclosures means that they can also contribute to the second-order goal of holding the governance system *itself* accountable – i.e. Kaminski's ideal of "second order accountability". As discussed, direct government regulation of social media recommendations is problematic from a fundamental rights perspective, since it applies opaque, technocratic methods to a highly contentious and politically sensitive field of governance. Even if multistakeholder perspectives from civil society or academia are incorporated in relevant oversight structures, such institutions run the risk of capture or

---

[145] James Grimmelman & Daniel Westreich (2017), 'Incomprehensible Discrimination', Cornell Legal Studies Research Paper No. 17-28. Available at:
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2950018
[146] Sandvig et al (2014), n. 55.
[147] Hargreaves et al (2018), 'Fairness in Online Social Network Timelines: Measurements, Models and Mechanism Design', *Performance Evaluation Review* 46:3.
[148] Ausloos (2019), n. 99.

bias. Releasing public information about content recommendation trends can help to critique such governance structures, and potentially even provide a starting point to identify more informal "jawboning" relationships between platform power and public power.

A similar argument about second-order accountability may also apply to other platform *users*, insofar as they also co-determine harmful outcomes in recommender systems. For instance, the discovery that certain harmful channels are being disproportionately recommended towards children could prompt intervention not only from platforms or from governments, but might also appeal to the responsibilities of the content provider in question. Public disclosures could ideally accelerate such media criticism by providing the necessary evidence of relevant recommendation trends. In sum, then, public records can assist civil society actors in holding not only platforms accountable but the governance system as a whole.

The benefits of public access pertain not only to civil society, but may also spill over to government oversight. Since regulatory auditing and other investigative powers can be slow and costly to perform, publicly available data can cut down on such costs and help agencies to perform high-level monitoring and more efficiently prioritise their in-depth investigative efforts. Perhaps more important, however, is the earlier point that regulatory enforcement commonly relies on knowledge sourced from third parties, e.g. through consultation responses, complaints, tips, and referrals, and scientific literature.[149] In this light, public data access can also redound to the benefit of government regulation, by helping third parties in monitoring social media recommendation systems, and referring cases to competent government agencies. For instance, content providers who depend on platforms to disseminate their content have incentives to monitor recommendations trends and check for potentially unlawful or anticompetitive patterns of discrimination. Public records could help them in such efforts, whereas government-focused transparency places the onus entirely on the government to do its own monitoring.

---

[149] Kaminski (2020, forthcoming), n. 66.

Of course, all of these potential benefits related to public access are still largely speculative, and their realisation depends on whether it is implemented effectively so as to offer meaningful and accessible information. The operational and technical challenges in designing such a regime are not to be underestimated, and further research is needed to pre-empt possible abuses. The above is simply intended to articulate the distinct benefits of public access, relative to more exclusive approaches currently seen in Europe policy. These benefits are particularly salient, it is submitted, in the politically sensitive context of media governance, where scepticism of both market and public ordering are uniquely strong and the demands of broad and inclusive second-order accountability are therefore particularly urgent.

4.2   Designing public access

A long line of transparency research has emphasised that transparency measures must be designed with the needs of their intended users in mind.[150] So what information, specifically, requires public access? This is a complex question, particularly since the potential userbase for public disclosures is necessarily undefined and open-ended, leaving it outside the scope of this paper to offer an exhaustive answer. Focusing on the needs of academia and civil society in particular, what follows is intended as exploratory, offering some starting points for further research and debate. In terms of format, public disclosures about recommender systems should include real-time, high-level, anonymized data access through public APIs and browser interfaces. In terms of content, public access should cover the documentation of recommendation outputs and their audiences; content-specific ranking decisions and other interventions by the operator's in recommendation system performance; and the organisational structures that control recommendation systems.

At the outset, an important starting point in terms of existing best practices is public research APIs. As discussed in Section 2.3, many social media platforms already offer some level of real-time public access through these systems, and public regulation can draw and

---

[150] Archon Fung et al (2007), *Full Disclosure: The Perils and Promise of Transparency*, Cambridge University Press.

build on this prior art. The functionality of these systems has been reduced significantly in recent years, nominally in response to privacy concerns resulting from the Cambridge Analytica scandal, but communications researchers argue that there is evidence of a disproportionate overreaction, and the pendulum has now swung too far back from openness to secrecy.[151] Binding public regulation could provide an impetus for (privacy-compliant) reform. To this end, policymakers can draw on expertise from communications science and adjacent fields, which command extensive experience with the design and usage of such public APIs.

As for the substance of public disclosures about content recommendations, one particularly salient aspect of their design which could be eligible for disclosure is content-specific ranking interventions. Platforms routinely intervene in recommender systems to alter specific outcomes, and such information could be eligible for disclosure. For instance, as mentioned in Section 2.2, Facebook currently partners with third-party fact-checkers to identify and downrank "false headlines" from untrustworthy news sources, and these fact-checkers publish explanations for each intervention they make. A more ambitious approach would register such decisions in a central platform repository, rather than dispersing them across various partnered websites. Ideally, such an approach would not only apply to third-party fact-checkers but to all human interventions in the algorithmic ranking system across the board, whether by platform workers or external partners. Such public records need not require full disclosure of the recommendation algorithm as a whole, as this could undermine service integrity and enable gaming of these systems by spammers and other malicious actors.[152]

An instructive comparison can be made between downranking and content removal decisions. For content removal decisions, platforms have declined content-level disclosure because the content at issue is by its very nature expected to be illegal or otherwise unsuitable for publication.[153] But this rationale does not apply when it comes to

---

[151] Section 2.3.
[152] E.g. Pasquale (2015), n. 3.
[153] q.v. Heidi Tworek & Paddy Leerssen (2019), *An Analysis of Germany's NetzDG Law*, Transatlantic High Level Working Group on Content Moderation Online and Freedom of Expression. Available at:

downranking decisions, since these are expressly intended for content that platforms do *not* wish to remove. In this light, there appears to be no compelling reason why these downranking decisions should not be made a matter of public record.

More broadly, it is worth investigating a best effort documentation requirement for other human-coded aspects of recommender algorithms. While many aspects of these algorithms are the product of complex machine-learning processes and therefore difficult to understand or explain even for their makers, other elements are human-coded and therefore easier to shed light on. One example is Facebook's Click-Gap initiative, which identifies low-quality based on the ratio of engagement on Facebook versus overall popularity across the web and thus serves to privilege more "mainstream", established media outlets.[154] It is to Facebook's credit that this update has been announced publicly.[155] But this is arguably the exception proving the more fundamental rule that conscious and explainable interventions are taking place, without any guarantee that these are necessarily disclosed to the public. How else have platforms intervened to curate their recommendations? Indeed, as discussed, YouTube boasts that it has made "hundreds" of changes in 2018 alone, and it is unclear what these entail.[156] A legal requirement that such interventions in the algorithm must be disclosed systematically would help to prevent any important omissions and underwrite the significance of platform disclosures. Of course, an important limitation is the risk of gaming in the algorithm, which may counsel against overly detailed specification of such changes: for instance if the specific keywords of an anti-spam blacklist were to be disclosed. Such defences may need to be evaluated on a case-by-case basis.

Public documentation of recommender systems need not focus exclusively on the *algorithm*. As discussed in Section 2.3, the algorithm as such cannot account fully for the effects and outcomes of recommender systems, as this requires reference to users' and

---

https://www.ivir.nl/publicaties/download/NetzDG_Tworek_Leerssen_April_2019.pdf . Keller & Leerssen (2020).
[154] See Section 2.2.
[155] Rosen (2019).
[156] YouTube (2019), n. 39.

their activity. As argued by Rieder, Matamoroz-Fernandez, and Coromina argue, this can best be approached through the study of recommendation *outcomes*, in terms of what content is recommended, and to whom.[157] At present, social media recommendations are scarcely documented for many important platforms. What are the most recommended pieces of content on YouTube or Instagram on a given day? In a given country? For a given age group? Some knowledge can be gleaned through independent observational methods, but, as discussed in Section 2.3, such methods face major operational challenges and necessarily produce incomplete and time-lagged datasets. Perhaps the most ambitious project in this space, Algotransparency.org, only covers YouTube recommendations made by 1000 selected channels and on a limited set of keywords.[158] While such methods have already led to important insights about social media recommendations,[159] far more comprehensive and systematic data could be published with the (regulated) cooperation of platforms themselves.[160] In essence, these output disclosures would serve to recreate, to some degree, the baseline publicity or "visibility" that was inherent in mass media content distribution, and has been lost through personalisation.[161] Even if the precise motives and decisions of our gatekeepers remain secret, at least the overall outcomes in terms of content distribution can then start to be observed.

Such public documentation of outputs would require strong safeguards against potential privacy harms. One important best practice is to limit disclosures to *publicly accessible* content—as opposed to private content such as personal messages. But this is no panacea: even though such an API would not technically expand access to content, since the content is already public, it would still make the content searchable and measurable in new ways

---

[157] Rieder, Matamoroz-Fernandez & Coromina (2017), n. 10.

[158] www.AlgoTransparency.org.

[159] E.g. Fisher & Taub (2019), n. 25. Paul Lewis (2018). 'Fiction is outperforming reality': how YouTube's algorithm distorts truth'. *The Guardian.* Available at:
https://www.theguardian.com/technology/2018/feb/02/how-youtubes-algorithm-distorts-truth;
Guillaume Chaslot (2016), 'YouTube's A.I. was divisive in the US presidential election', *Medium*.
Available at: https://medium.com/the-graph/youtubes-ai-is-neutral-towards-clicks-but-is-biased-towards-people-and-ideas-3a2f643dea9a

[160] Cobbe & Singh (2019), n. 2.

[161] Section 2.3.

for third parties, which may raise privacy issues of its own.[162] In addition, therefore, public access can also apply a *de minimis* rule: only content above a certain threshold of popularity could be included. Such a rule, already commonly seen in public research APIs, would limit the scope to the most important and visible content, and protect more sensitive activity from excessive monitoring. Potentially, content below this threshold could still be described with certain metadata, such as keywords, format, or language, to provide at least some basic insight into content flows without threatening the underlying content.

In designing public access regimes and their privacy safeguards, a relevant precedent is the experience with political advertising archives. Since 2018, major social media platforms have started developing such public archives to document political ads sold on their services.[163] Like the output documentations discussed in this paper, ad archives are similarly concerned with accountability in the algorithmically personalised distribution of content—albeit in the particular context of advertising content. In ad archives, the data is disclosed through searchable web interfaces as well as through APIs, and audience data are anonymised and aggregated to avoid user privacy concerns. It should be noted that present self-regulatory implementations have been criticised extensively; for instance, researchers from Mozilla concluding that the API was so bug-ridden as to be effectively unusable—a strong argument for regulation in this space.[164] Nevertheless, whilst their research utility is necessarily limited for understanding deeper questions about algorithmic sorting and bias, platform ad archives have already started to see regular use in real-time media monitoring and election coverage.[165] In this light, the experience with ad archives is instructive in two ways. First, it warns against an overreliance on self-regulation, given the

---

[162] Michael Zimmer (2010), '"But the data is already public": on the ethics of research on Facebook', *Ethics in Information Technology,* 12:4.

[163] Paddy Leerssen et al (2019), 'Platform ad archives: promises and pitfalls', *Internet Policy Review* 8:4.

[164] Matthew Rosenberg (2019). 'Ad Tool Facebook Built to Fight Disinformation Doesn't Work as Advertised', *The New York Times.* Available at:
https://www.nytimes.com/2019/07/25/technology/facebook-ad-library.html

[165] See, for instance: Julia Carry Wong (2020), 'One year inside Trump's monumental Facebook campaign'. *The Guardian.* Available at: https://www.theguardian.com/us-news/2020/jan/28/donald-trump-facebook-ad-campaign-2020-election?CMP=Share_iOSApp_Other

critical failures of voluntary initiatives in this space. Second, despite their inadequate implementation in practice, ad archives do provide a basic conceptual blueprint for public transparency in algorithmic content distribution: real-time, anonymised, output-focused, and accessible to all.

A final point of attention for public access is the organisations behind recommender systems. Information about these organisations is highly relevant to understand how gatekeeping decisions are made, and better outcomes can be ensured. Relevant issues in the context of recommender systems could include the location, demographic background, training, reward schemes, authorisations, and management systems in place for relevant workers. Comparable rules about organisational transparency can already be found in Germany's *Netzwerkdurchsetzungsgesetz*, which includes public documentation requirements for staffing and training of content removal operations related to this law.[166] Professional standards for transparency in journalistic organisations can also serve as a template.

4.3   Regulating public access

Like most forms of platform transparency, public access will require a binding legal basis in order to be effective. As discussed in Section 2.3, platforms have a poor track record in their voluntary transparency reforms, and even when cooperating in earnest may still be dogged by questions of credibility and independence.[167] Binding transparency obligations can help to address these concerns, and avoid a situation in which "only approved questions get answered".[168] In addition, public regulation of transparency can help to offset legal restrictions on data disclosure, e.g. by providing relevant exemptions under intellectual property law, or processing grounds under data protection law.

---

[166] Tworek & Leerssen (2019), n. 154. Wagner et al (2020), n. 71.
[167] Section 2.3.
[168] Brother Ali (2007), 'Uncle Sam Goddamn'. *The Undisputed Truth* [CD]. Minneapolis: Rhymesayers Entertainment.

Another advantage of binding regulation is that it may remedy the current precarity of automated research access, discussed in Section 2.3. Bernhard Rieder & Jeannette Hoffman argue that the goal in platform governance should be to "transpos[e] local experiments into more robust practices able to guarantee continuity and accumulation", leading to "structured interfaces between platforms and society".[169] Relevant to this endeavour is Fung et al's research on the "sustainability" of transparency measures, which recommends that they improve in scope, accuracy, and use over time.[170] With this in mind, it is advisable for a regulatory effort to start with a relatively modest scale, perhaps as a pilot study or experiment, and then gradually expand in respond to feedback from early users.[171] Other elements of sustainable transparency highlighted by Fung et al include effective enforcement of applicable rules, and the strengthening of potential user groups such as civil society organisations.[172]

Due to platform dominance dynamics,[173] size-based regulation is appropriate; targeting the most influential platforms addresses the major sources of risk while avoiding unnecessary or disproportionate burdens on smaller services. For instance, platform size could be defined based on revenue, user count, view count, or some combination of these metrics. Similar size-based regulation is already common in recent proposals for transparency in social media platforms, such as the EU Code of Practice, the US Honest Ads Act and Germany's *Netzwerkdurchsetzungsgesetz.*

Given the complexity of designing privacy-compliant disclosure standards, rules for public access will be difficult to codify exhaustively in one-size-fits-all legislation. Not only are social media recommendations technically complex, they are also heterogeneous; each has

---

[169] Bernhard Rieder & Jeannette Hoffman (forthcoming), 'Beyond Transparency: Toward Platform Observability'. On file with author.

[170] Archon Fung et al (2007), *Full Disclosure: The Perils and Promise of Transparency*, Cambridge University Press.

[171] Jef Ausloos, Paddy Leerssen & Pim ten Thije, *Operationalizing Research Access in Platform Governance: What To Learn From Other Industries?* Available at: https://algorithmwatch.org/en/governing-platforms-ivir-study-june-2020/#study (arguing for a pilot-driven, iterative approach to transparency regulation.)

[172] Fung et al (2007), n. 171, describe civil society organizations as 'user intermediaries', who mediate between the disclosing party and the ultimate end users of information.

[173] Barwise & Watkins (2018), n. 97.

unique features (types of posts and formats, engagement metrics, et cetera) which may require unique forms of documentation and privacy safeguards. In response, disclosure standards in legislation must remain broad, to be specified on a case-by-case basis by an authorised regulator. Of course, such a body could also be instrumental in achieving other transparency goals in social media governance besides public access, such as those related to individual user notices, regulatory enforcement and exclusive research access frameworks.

An ongoing challenge regarding transparency regulation is finding an appropriate regulatory body to enforce these rules. Few national systems have developed agencies equipped to regulate social media, and leaving Member States to each develop their own institutional capacities risks not only a duplication of efforts but also the risk of regulatory fragmentation and (potentially conflicting) standards. Transparency measures, like most information products, tend to benefit from economies of scale, which support the case for uniform regulation at EU level. And yet, a (social) media regulator does not yet exist at the EU level. Indeed, member states have historically resisted the creation of a EU media regulator given the cultural and political sensitivities in this space.[174]

Whilst developing a definitive division of competences is outside the scope of this paper, it is worth emphasising that the regulation of transparency may in theory be separated from substantive media policy. Under such an approach, the EU could put its full force behind ensuring access to information, whilst leaving national entities to make use of this data for their various regulatory efforts and thus to realise the *substance* of domestic media policy.[175] This aligns with other proposals to institute government regulators focused on transparency. For instance, Ben Wagner and Lubos Kuklis envisage a "single European institution which could act as an auditing intermediary to ensure that the data provided to regulators by social media companies are accurate".[176] Their proposal focuses on

---

[174] Klimkiewicz (2009), n. 115.
[175] Klimkiewicz (2009), n. 115.
[176] Ben Wagner & Lubos Kuklis (2019), 'Disinformation, data verification and social media'. *Media@LSE.* Available at: https://blogs.lse.ac.uk/medialse/2020/01/07/disinformation-data-verification-and-social-media/
https://blogs.lse.ac.uk/medialse/2020/01/07/disinformation-data-verification-and-social-media/

transparency towards other public regulators, such as data protection and competition authorities, but a similar vision could also apply to public access and its use by a range of governance stakeholders in government and civil society.[177] The ideas about public access regulation outlined in this paper should be considered alongside such broader debates about the need for dedicated regulatory structures for transparency for platforms in general, and social media in particular.

5   Conclusion

These are decisive times for the regulation of social media content recommendations. As the "techlash" moves from opinion pages to public policy, and attempts at regulation begin in earnest, we see a variety of attempts to make social media platforms more transparent and accountable in their content recommendations. A governance landscape is emerging in which users, governments and civil society all have a role to play in holding these systems accountable, and realising public values in our content feeds. Transparency rules are developing accordingly, with each stakeholder group being associated with its own types of disclosures. As recurring themes in ongoing policy, this paper has identified notices and disclaimers, government auditing, and data access partnerships.

A central component in ongoing efforts is the enabling of independent oversight by academia and civil society. This is laudable given the particular sensitivity of recommender governance from the perspective of democracy and fundamental rights. Yet this paper has cautioned against efforts which pursue transparency towards academia and civil society exclusively through institutionalised systems of privileged data access. Whilst such privileged access regimes have important advantages in enabling in-depth scholarly research, there may be low-hanging fruit of non-sensitive data that could find far wider uptake if made public without restriction. This paper has articulated how real time, high level public access has distinct advantages for accountability in this space. A robust system of public access not only allows for wider uptake and greater impact, but is essential to

---

[177] See, for instance, French Secretary for Digital Affairs (2019), n. 127.

make the technocratic, expert-driven institutions of recommendation governance accountable to scrutiny and contestation by broader publics and interest groups.

This paper has also provided some starting points for the design and regulation of such public access. Overall, it suggests a reorientation from "algorithms" as objects of transparency towards a broader inquiry into the sociotechnical dynamics of recommender systems. To this end, fruitful avenues for public access include content-level detail on downranking decisions and other manual interventions in the recommender system, as well as publicly searchable documentation of recommendation outputs for the most popular content. More generally, policymakers should explore existing best practices in the design and regulation of public research APIs. Given the complexity of these issues, the most promising way to regulate this would be through broad legislative standards, specified and enforced by an authorised regulator. This approach resonates with recent proposals in academia and government to install a dedicated transparency regulator for online platforms.

On a final note: This paper's discussion of transparency has hewed closely to the prevailing vision in European media governance of social media platforms as regulated oligopolists, whose dominance as online speech infrastructure is not to be replaced or contested but rather to be made more transparent and accountable to public interest considerations. It remains to be seen, given the vast power and complexity of these services and the sensitivity of the data they process, whether such a vision can be realised. We may well come to conclude that for-profit stewardship of these influential and opaque systems simply creates unacceptable and unmanageable risks to democracy; that meaningful transparency in these circumstances is a false hope—much less accountability—and that instead more fundamental changes to ownership or business models may be necessary, such as switching to cooperatively owned or publicly owned social media services.[178] But

---

[178] Evgeny Morozov (2019), 'Digital Socialism? The Calculation Debate in the Age of Big Data'. *New Left Review* 116. Trebor Scholz (2016). *Platform Cooperativism: Challenging the Corporate Sharing Economy*. Rosa Luxemburg Stiftung. Ethan Zuckerman (2020), 'The Case for Digital Public Infrastructure'. Knight First Amendment Institute. Available at: https://knightcolumbia.org/content/the-case-for-digital-public-infrastructure. Mariana Mazzucato

even for these more radical visions of online media governance, the arguments discussed in this paper may still hold some relevance: what will likely remain are the importance of broad and inclusive scrutiny of algorithmic gatekeeping, and the distinct benefits of publicly accessible information to that end.

6　　Acknowledgements

(2018), 'Let's Make Data Into A Public Good'. *MIT Policy Review.* Available at: https://www.technologyreview.com/s/611489/lets-make-private-data-into-a-public-good/