



RESEARCH
ARTICLE



OPEN
ACCESS



PEER
REVIEWED

Towards platform observability

Bernhard Rieder *University of Amsterdam* B.Rieder@uva.nl

Jeanette Hofmann *Berlin Social Science Center (WZB)*

DOI: <https://doi.org/10.14763/2020.4.1535>

Published: 18 December 2020

Received: 4 August 2020 **Accepted:** 28 October 2020

Competing Interests: The author has declared that no competing interests exist that have influenced the text.

Licence: This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 License (Germany) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. <https://creativecommons.org/licenses/by/3.0/de/deed.en>
Copyright remains with the author(s).

Citation: Rieder, B. & Hofmann, J. (2020). Towards platform observability. *Internet Policy Review*, 9(4). <https://doi.org/10.14763/2020.4.1535>

Keywords: Platforms, Algorithms, Transparency, Observability, Regulation

Abstract: The growing power of digital platforms raises the question of democratic control or at least containment. In light of the transforming impact of platforms on markets, the public sphere, elections, and employment conditions, governments, and civil society alike are demanding more transparency and accountability. Shedding light on the principles and practices of algorithmic ordering promises to limit the power of platforms by subjecting their hidden operations to regulatory inspection. This article questions the popular image of an openable 'black box'. Based on a critical reflection on transparency as a panacea for curtailing platform power, we propose the concept of observability to deal more systematically with the problem of studying complex algorithmic systems. We set out three broad principles as regulatory guidelines for making platforms more accountable. These principles concern the normative and analytical scope, the empirical and temporal dimension, and the necessary capacities for learning and knowledge generation.

1. Introduction

Platforms are large-scale infrastructures specialised in facilitating interaction and exchange among independent actors. Whether understood economically as two- or multi-sided markets (Langley & Leyshon, 2017) or with an eye on online media as services that ‘host, organize, and circulate users’ shared content or social interactions’ (Gillespie, 2018, p. 18), platforms have not only become highly visible and valuable companies but also raise important social challenges. While intermediaries have in one form or another existed for millennia, contemporary platforms are relying on digital technologies in (at least) two fundamental ways. First, platforms ‘capture’ (Agre, 1994) activities by channelling them through designed functionalities, interfaces, and data structures. Uber, for example, matches riders with drivers in physical space, handles payment, and enforces ‘good behaviour’ through an extensive review system covering both parties. This infrastructural capture means that a wide variety of data can be generated from user activity, including transactions, clickstreams, textual expressions, and sensor data such as location or movement speed. Second, the available data and large numbers of users make algorithmic matching highly attractive: ranking, filtering, and recommending have become central techniques for facilitating the ‘right’ connections, whether between consumers and products, users and contents, or between people seeking interaction, friendship, or love.

Digital platforms host social exchange in ways that Lawrence Lessig (1999) summarised under the famous slogan ‘code is law’, which holds that technical means take part in regulating conduct and shaping outcomes. The combination of infrastructural capture and algorithmic matching results in forms of socio-technical *ordering* that make platforms particularly powerful. As Zuboff (2019, p. 15) discusses under the term *surveillance capitalism*, the tight integration of data collection and targeted ‘intervention’ has produced ‘a market form that is unimaginable outside the digital milieu’. The rising power of platforms poses the question of what kind of accountability is necessary to understand these processes and their consequences in more detail. Matching algorithms, in particular, represent ordering mechanisms that do not follow the same logic as traditional decision-making, leading to considerable uncertainty concerning their inner workings, performativities, and broader social effects.

So far, most regulatory approaches to tackling these questions seek to create accountability by ‘opening the black box’ of algorithmic decision-making. A recent EU regulation on fairness in platform-to-business relations, for example, proposes *transparency* as its principal means.¹ The public debate about the upcoming EU

Digital Services Act indeed shows that calls for transparency of algorithmic power have gained support across parliamentary factions and stakeholder groups.² The ‘Filter Bubble Transparency Act’—a US legislative proposal that seeks to protect users from being ‘manipulated by algorithms driven by user-specific data’ - focuses more specifically on platforms as media, but again relies on transparency as guiding principle.³ The German *Medienstaatsvertrag* (‘State Media Treaty’), which has recently been ratified by all state parliaments, explicitly requires platform operators to divulge criteria for ranking, recommendation, and personalisation ‘in a form that is easily perceivable, directly reachable, and permanently available’.⁴ This widespread demand for disclosure and explanation articulates not only justified concerns about the opacity of platforms but also testifies to the glaring lack of information on their conduct and its social, political, and economic repercussions.

In this paper, we likewise take up the challenge posed by platform opacity from the angle of accountability but seek to probe the conceptual and practical limitations of these transparency-led approaches to platform regulation. Echoing the critical literature on transparency as a policy panacea (e.g., Etzioni, 2010; Ananny & Crawford, 2018), we propose the concept of *observability* as a more pragmatic way of thinking about the means and strategies necessary to hold platforms accountable. While transparency and observability are often used synonymously (e.g. August & Osrecki, 2019), we would like to highlight their semantic differences. Unlike transparency, which nominally describes a state that may exist or not, observability emphasises the *conditions for the practice of observing* in a given domain. These conditions may facilitate or hamper modes of observing and impact the capacity to generate external insights. Hence, while the image of the black box more or less skips the practicalities involved in opening it, the term observability intends to draw attention to and problematise the process dimension inherent to transparency as a regulatory tool.

While observability incorporates similar regulatory goals to transparency, it also deviates in important respects, most importantly by understanding accountability as a complex, dynamic ‘social relation’ (Bovens, 2007, p. 450), which is embedded in a specific material setting. The goal is not to exchange one concept for the oth-

1. <https://eur-lex.europa.eu/eli/reg/2019/1150/oj>

2. See, for example, the response by AlgorithmWatch and other signatories to the European Commission’s planned Digital Services Act: <https://algorithmwatch.org/en/submission-digital-services-act-dsa/>.

3. <https://www.congress.gov/bill/116th-congress/senate-bill/2763/all-info>

4. https://www.rlp.de/fileadmin/rlp-stk/pdf-Dateien/Medienpolitik/ModStV_MStV_und_JM-StV_2019-12-05_MPK.pdf

er but to sharpen our view for the specificities of platform power. At the risk of stating the obvious, regulatory oversight needs to take into account the material quality of the objects under investigation. Inspecting the inner workings of a machine learning system differs in important ways from audits in accounting or the supervision of financial markets. Rather than nailing down ‘the algorithm’, understood as a singular decision mechanism, the concept of observability seeks to address the conditions, means, and processes of knowledge production about large-scale socio-technical systems. In the everyday life of platforms, complex technologies, business practices, and user appropriations are intersecting in often unexpected ways. These platform dynamics result in massive information asymmetries that affect stakeholder groups as well as societies at large. Regulatory proposals need to take a broader view to live up to these challenges.

Our argument proceeds in three steps. In the next section, we retrace some of the main problems and limitations of transparency, paying specific attention to technical complexity. The third section then discusses the main principles guiding the observability concept and provides concrete examples and directions for further discussion. We conclude by arguing for a policy approach to promoting observability, emphasising that institutional audacity and innovation are needed to tackle the challenges raised by digital platforms.

2. Limitations to transparency

Much of the debate around our insufficient understanding of platforms and their use of complex algorithmic techniques to modulate users’ experience has centred on the metaphor of a ‘black box’. Although Frank Pasquale, whose *Black Box Society* (2015) has popularised the term beyond academia, prefers the broader concept of *intelligibility*, the talk of black boxes is often accompanied by demands for *transparency*. The regulatory proposals mentioned above are largely organised around mechanisms such as explanations, disclosures, and—more rarely—audits⁵ that would bring the inner workings of the machine to light and thereby establish some form of control. But these calls for transparency as a remedy against unchecked platform power encounter two sets of problems. First, the dominant understanding of transparency as information disclosure faces important limitations. Second, the object under scrutiny itself poses problems. Platforms are marked by opacity and complexity, which effectively challenges the idea of a black box whose lid can be

5. The ACM’s Statement on Algorithmic Transparency and Accountability (https://www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf), for example, explicitly mentions ‘auditability’ as a desirable principle.

lifted to look inside. This section discusses both of these issues in turn.

2.1. Accountability as mediated process

Transparency has a long tradition as a ‘light form’ (Etzioni, 2010) of regulation. It gained new popularity in the 1970s as a neoliberal governance method, promising better control of organisational behaviour through inspection (August & Osrecki, 2019). Transparency is seen as an essential means of oversight and of holding commercial and public entities to account: only if powerful organisations reveal relevant information about their actions are we able to assess their performance. This understanding of transparency implies a number of taken for granted assumptions, which link information disclosure to visibility, visibility to insight, and insight to effective regulatory judgement (Ananny & Crawford, 2018, p. 974). According to this view, transparency is able to reveal the truth by reflecting the internal reality of an organisation (Albu & Flyverbom, 2019, p. 9) and thereby creating ‘representations that are more intrinsically true than others’ (Ananny & Crawford, 2018, p. 975). Making the opaque and hidden visible, creates truth and truth enables control, which serves as a ‘disinfectant’ (Brandeis, 1913, p. 10) capable of eliminating malicious conduct. Transparency is considered crucial for the accountability of politics because seeing, just as in the physical world, is equated with knowing: ‘what is seen is largely what is happening’, as Ezrahi (1992, p. 366) summarises this view. These assumptions also inform current considerations on platform regulation.

However, recent research on transparency has shown that transparency does more and different things than shedding light on what is hidden. The visibility of an entity and its procedures is not simply a disclosure of pre-existing facts, but a process that implies its own perspective. While transparency requirements expect ‘to align the behavior of the observed with the general interest of the observers’, empirical studies found that ‘transparency practices do not simply make organizations observable, but actively change them’ (August & Osrecki, 2019, p. 16). As Flyverbom (2016, p. 15) puts it, ‘transparency reconfigures - rather than reproduces - its objects and subjects’. The oversight devices used to generate visibility shape what we get to see (Ezrahi, 1992; Flyverbom, 2016), which puts into question the idea of direct, unmediated access to reality if only the disclosed information is *accurate*.

From a social science perspective, transparency should not be regarded as a state or a ‘thing’ but as the practice ‘of deciding what to make present (i.e. public and transparent) and what to make absent’ (Rowland & Passoth, 2015, p. 140). Creating visibility and insights as part of regulatory oversight consists of specific proce-

dures, which involve choices about what specifically should be exposed and how, what is relevant and what can be neglected, which elements should be shown to whom and, not least, how the visible aspects should be interpreted (Power, 1997). In their critique of transparency-led approaches to algorithmic accountability, Ananny & Crawford (2018) moreover argue that there is a distinct lack of sensitivity for fundamental power imbalances, strategic occlusions, and false binaries between secrecy and openness, as well as a broad adherence to neoliberal models of individual agency.

In light of these criticisms, it may not come as a surprise that regulatory transparency obligations often fall short of their goals and create significant side-effects instead. Among the most common unintended outcomes are bureaucratisation, generalised distrust, and various forms of ‘window dressing’ designed to hide what is supposed to be exposed to external review. Informal organisational practices emerge and coexist with official reports, accounts, and presentations (August & Orecki, 2019, p. 21). While the critical literature on regulatory failures of transparency obligations is increasing, these insights have yet to have an impact on regulatory thinking. Most regulatory proposals resort to traditional ideas of external control through transparency and frame transparency as a straightforward process of disclosure. As a result, they are missing the mark on the complex and conflictual task of creating *meaningful* understanding that can serve as an effective check on platform power.

Taken together, a social science perspective on this key ideal of regulation suggests that making platforms accountable requires a critical engagement with the achievements and shortcomings of transparency. It needs to take on board efforts to combine different forms of evidence, and above all, to become attentive to the selective and mediated character of knowledge-building. Similar to the flawed logic of ‘notice and consent’ in the area of privacy protection, which holds that informing individuals on the purposes of data collection allows them to exercise their rights, a superficial understanding of transparency in the area of platform regulation risks producing ineffective results (see Obar, 2020; Yeung, 2017).

2.2. Opacity, complexity, fragmentation

A second set of complications for transparency concerns algorithms and platforms as the actual objects of scrutiny. Large-scale technical systems, in particular those incorporating complex algorithmic decision-making processes, pose severe challenges for assessing their inner workings and social effects. One obvious reason for this is indeed their opacity. As Burrell (2016, p. 2) argues, opacity may stem

from secrecy practices, lack of expertise in reading code, and the increasing ‘mismatch between mathematical optimization in high-dimensionality characteristic of machine learning and the demands of human-scale reasoning’. The last point in particular introduces significant challenges to transparency understood as information disclosure or audit. Even if decision procedures behind automated match-making can sometimes still be meticulously specified, platforms nowadays mainly deploy statistical learning techniques. These techniques develop decision models inductively and ‘learn programs from data’ (Domingos, 2012, p. 81), based on an arrangement between data, feedback, and a given purpose (see Rieder, 2020).

In the canonical example of spam filtering, users label incoming emails as spam or not spam. Learning consists in associating each word in these messages with these two categories or ‘target variables’. Since every word contributes to the final decision to mark an incoming message as spam or not spam, the process cannot be easily traced back to singular factors. Too many variables come into play, and these algorithms are therefore not ‘legible’ in the same way as more tangible regulatory objects. With regard to regulatory oversight, this means that transparency in the sense of reconstructing the procedure of algorithmic decision making ‘is unlikely to lead to an informative outcome’, as Koene et al. (2019, p. II) conclude. Audits are unable to find out ‘what the algorithm knows because the algorithm knows only about inexpressible commonalities in millions of pieces of training data’ (Dourish, 2016, p. 7). There is a large gulf between the disclosure of ‘fundamental criteria’ mandated by regulatory proposals like the *Medienstaatsvertrag* and the technical complexities at hand.

Even if regulators were given access to data centres and source code, the process of sense-making would not be straightforward. Reading the gist of an algorithm from complex code may run into difficulties, even if no machine learning is involved. As Dourish (2016) shows, the presence of different programming languages and execution environments adds further complications, and so do the many subsystems and modules that concrete programmes often draw on. Algorithmic decision procedures ‘may not happen all in one place’ (Dourish, 2016, p. 4) but can be distributed over many different locations in a large programme or computer network. In the case of online advertising, for example, the placement of a single ad may entail a whole cascade of real-time auctions, each drawing on different algorithms and data points, each adding something to the final outcome. The result is a continuously evolving metastable arrangement. Thus, time becomes a crucial analytical factor, causing considerable difficulties for the ‘snapshot logic’ underlying most audit proposals.

For these reasons, algorithms turn out to be difficult to locate. In his ethnographic study of a recommender system, Seaver (2017) observes that even in small companies it can be a challenge for staff members to explain where exactly ‘the algorithm’ is. As Bogost (2015) quips, ‘[c]oncepts like “algorithm” have become sloppy shorthands, slang terms for the act of mistaking multipart complex systems for simple, singular ones’. What is referred to as ‘algorithm’, i.e. the actual matchmaking technique, may thus only be a small component in a much larger system that includes other various instances of ordering, ranging from data modelling to user-facing interfaces and functions that inform and define what users can see and do. YouTube, for example, not only fills its recommendation pipeline with a broad array of signals generated from the activities of billions of users but actually uses two different deep learning models for ‘candidate generation’ (the selection of hundreds of potential videos from the full corpus) and ‘ranking’ (the selection and ordering of actual recommendations from the candidate list) (see Covington et al., 2016). The fuzzy, dynamic, and distributed materiality of contemporary computing technologies and data sets means that algorithmic accountability is harder to put into practice than the call for transparency suggests. Regulatory proposals such as disclosures, audits, or certification procedures seeking to establish effective control over their functionality and effects assume properties that algorithmic systems may often not meet. Suffice to say that technical complexity also facilitates the attempts at dissimulation and ‘window dressing’ mentioned above.

Yet, as if this was not difficult enough, our understanding of *platform* accountability should extend beyond oversight of algorithms and platform conduct to be meaningful. The ordering power of platforms also encompasses shared or distributed accomplishments (see Suchman, 2007) to which platforms, users and content providers each contribute in specific ways. As Rahwan et al. (2019, p. 477) argue, machine behaviour ‘cannot be fully understood without the integrated study of algorithms and the social environments in which algorithms operate’. The actions of users, for example, provide the data that shape algorithmic models and decisions as part of machine learning systems. In the same vein, platform *behaviour* cannot be reduced to platform *conduct*, that is, to the policies and design decisions put in place by operators. It must include the evolving interactions between changing social practices and technical adjustments, which may, in turn, be countered by user appropriations. As use practices change, algorithmic decision models change as well. Platform companies are therefore neither fully in control of actual outcomes, nor fully aware what is happening within their systems.

Finally, the effects of platforms can only be sufficiently addressed if we consider

what is being ordered. For example, ranking principles considered beneficial in one culture domain, e.g. music recommendation, may have troubling implications in another, e.g. the circulation of political content. Accountability thus has to consider what is made available on platforms and how ordering mechanisms interact with or shape the content and its visibility. This again requires a broader view than what algorithm audits or broad technical disclosures are able to provide.

Taken together, research on the properties of algorithms and algorithmic systems suggests that regulatory proposals such as ‘opening the black box’ through transparency, audit, or explainability requirements reflect an insufficient understanding of algorithms and the platform architectures they enable. Algorithms can neither be studied nor regulated as single, clear-cut, and stable entities. Rather, their behaviour and effects result from assemblage-like contexts whose components are not only spatially and functionally distributed but also subject to continuous change, which is partly driven by users or markets facilitated by platforms. Given the ephemeral character of algorithms on the one side and the enormous generative and performative power of algorithmic systems on the other, the question arises what concepts, strategies, and concrete tools might help us to comprehend their logics and to establish effective political oversight. Such an approach needs to take on board the critique of transparency as a regulatory tool and consider accountability as a continuous interaction and learning process rather than periodical undertakings. It should recognise that the legibility of algorithmic systems significantly differs from that of other objects or areas of regulation; and it should take into account that any form of review is not only selective but also shapes the object under investigation. Thus, the debate on platform regulation needs to become reflexive with regard to the specific materiality of the regulatory field and the constitutive effects of studying it.

3. Principles of observability

This section seeks to flesh out an understanding of *observability* as a step toward tackling the problems platform accountability currently faces. While the term is regularly used in the literature on transparency (e.g., Bernstein, 2012; Albu & Flyverbom, 2015; August & Osrecki, 2019), we seek to calibrate it to our specific goals: the challenges raised by platforms as regulatory structures need to be addressed more broadly, beginning with the question of how we can assess what is happening within large-scale, transnational environments that heavily rely on technology as a mode of governance. Who gets treated how on large online platforms, how are connections between participants made and structured, what are

the outcomes, and—crucially—who can or should be able to make such assessments? Rather than a binary between transparency and opacity, the question is how to foster the capacity to produce knowledge about platforms and ‘platform life’ in constructive ways. The increasingly technological nature of our societies requires not just penalties for law infringements, but a deeper and well-informed public conversation about the role of digital platforms. This includes attention to the larger impacts of the new kinds of ordering outlined above, as well as a sensitivity for the ideological uses of transparency, which may serve ‘as a tool to fight off the regulations opposed by various business groups and politicians from conservative parties’ (Etzioni, 2010, p. 2). We therefore position observability as an explicit means of, not an alternative to regulation. As van Dijck et al. (2018, p. 158) underline, ‘[r]egulatory fixes require detailed insights into how technology and business models work, how intricate platform mechanisms are deployed in relation to user practices, and how they impact social activities’. Our concept of observability thus seeks to propose concrete actions for how to produce these insights. While some of the more concrete strategies we discuss may come out of self-regulation efforts, effective and robust observability clearly requires a regulatory framework and institutional support. In what follows, we outline three principles that inform the concrete conceptual and practical directions observability seeks to emphasise.

3.1. Expand the normative and analytical horizon

The first principle concerns the research perspective on platforms and argues that a broader focus is needed. This focus takes into consideration how digital platforms affect societies in general, ranging from everyday intimacy to economic and labour relations, cultural production, and democratic life. Given that *platformisation* transforms not only specific markets but ‘has started to uproot the infrastructural, organizational design of societies’ (van Dijck, 2020, p. 2), it seems crucial to develop knowledge capacities beyond critical algorithm studies and include platform conduct, behaviour, and effects across relevant social domains in our agendas. As Powles and Nissenbaum (2018) have recently argued for artificial intelligence systems, limiting our focus to the important yet narrow problems of fairness and biases means that ‘vast zones of contest and imagination are relinquished’, among them the question whether the massive efforts in data collection underlying contemporary platform businesses are acceptable in the first place. The ability to say no and prohibit the deployment of certain technologies such as political micro-targeting of voters or face recognition requires robust empirical and normative evidence on its harm for democracies.

While investigations into misinformation and election tampering are important,

there are other long-term challenges waiting to be addressed. Recent studies on surveillance capitalism (Zuboff, 2019), digital capitalism (Staab, 2019), informational capitalism (Cohen, 2019), the platform society (van Dijck et al., 2018), or the ‘dataist state’ (Fourcade & Gordon, 2020) aim to capture and make sense of the ongoing structural changes of societies and economies, including the power shifts these imply. EU commissioner Vestager recently evoked Michel Foucault’s notion of *biopower* when addressing novel data-based techniques of classifying, sorting, and governing (Stolton, 2019). While the term addresses a set of political technologies that emerged in the 19th century to manage the behaviour of populations by means of specific regimes of knowledge and power, digital platforms’ considerable reach and fine-grained ‘capture’ (Agre, 1994) of everyday activities invites comparison. The deep political and social repercussions these conceptual frames highlight require broader forms of *social accountability* (Bovens, 2007) than disclosures or audits are able to provide.

How can researchers, regulators, and civil society expand their capacity to study, reflect and act on these developments? The concept of observability starts from the recognition of a growing information asymmetry between platform companies, a few data brokers, and everyone else. The resulting data monopoly deprives society of a crucial resource for producing knowledge about itself. The expanding data sets on vast numbers of people and transactions bear the potential for privileged insights into societies’ texture, even if platforms tend to use them only for operational purposes.

AirBnB’s impact on urban development, Uber’s role in transforming transportation, Amazon’s sway over retail, or Facebook and Twitter’s outsized influence on the public sphere cannot be assessed without access to relevant information. It is symptomatic that companies refuse access to the data necessary for in-depth, independent studies and then use the lack of in-depth, independent studies as evidence for lack of harm. New modes of domination are unfolding as part of analytics-driven business models and the unprecedented information asymmetries they bring about. Powles and Nissenbaum (2019) therefore argue that we need ‘genuine accountability mechanisms, external to companies and accessible to populations’. An essential condition and experimental construction site for such accountability mechanisms would be the institutionalisation of reliable information interfaces between digital platforms and society—with a broad mandate to focus on the *public interest*.

We propose the concept of public interest as a normative reference for assessing platform behaviour and regulatory goals. However, public interest is neither well

defined nor without alternatives.⁶ We prefer *public interest* over the closely related *common good* because the former refers to an internationally established mandate in media regulation and could thus inform the formulation of specific requirements or ‘public interest obligations’ for platforms as well (Napoli, 2015, p.4). Furthermore, the concept speaks to our specific concern with matters of governance of platform life. The use of public interest spans different disciplinary and regulatory contexts, and it is open to flexible interpretation. Yet, the often-criticised vagueness of the concept has the advantage of accommodating the broad range of existing platforms. As a normative framework it can be used to critically assess the design of multiple-sided markets as much as the impact of digital intermediaries on the public sphere. Approaches to defining and operationalising public interest depend on the context. In economic theory, public interest is suspected of functioning as a ‘weapon’ for justifying regulatory intervention into markets for the purpose of enhancing social welfare (Morgan & Yeung, 2007). Correcting failing markets constitutes a minimalist interpretation of public interest, however. In politics, public interest is associated with more diverse social goals, among them social justice, non-discrimination, and access to social welfare; or more generally the redistribution of resources and the maintenance of public infrastructures. With regard to the public sphere and the media sector, public interest refers to protecting human rights such as freedom of information and freedom of expression, fostering cultural and political diversity, and not least sustaining the conditions for democratic will formation through high quality news production and dissemination (Napoli, 2015).

What these different understandings of public interest have in common is a focus on both procedural and substantial aspects. Obviously, public interest as a frame of reference for assessing and regulating digital platforms is not a given. Rather, the meaning and principles of public interest have to be constantly negotiated and reinterpreted. As van Dijck (2020, p. 3) reminds us, such battles over common interest do not take place in a vacuum, they are ‘historically anchored in institutions or sectors’ and ‘after extensive deliberation’ become codified in more or less formal norms. From a procedural point of view, public interest can also be defined as a practice, which has to meet standards of due process such as inclusiveness, transparency, fairness, and right to recourse (Mattli & Woods, 2009, p. 15). In terms of substance, the notion of public interest clearly privileges the collective common

6. For a discussion of the intricate history of ideas behind the concepts of the common good and public interest in the anglo-american realm and a definition of the latter see Douglass (1980, p. 114): ‘the public interest would come to mean what is really good for the whole people. And in a democratic society, this would mean what is really good for the whole people as interpreted by the people.’

welfare over that of individuals or private commercial entities. In this respect, it entails a departure from the neoliberal focus on individual liberty toward collective freedoms. Thereby it also extends the space of policy options beyond ‘notice and consent’ to more far-reaching regulatory interventions (Yeung, 2017, p. 15). We see similar conceptual adjustments toward public interest in other areas such as the discourse on data protection. As Parsons (2015, p. 6) argues, it is necessary to recognise ‘the co-original nature of [...] private and public autonomy’ to understand that mass surveillance is not merely violating citizens’ individual rights, but ‘erodes the integrity of democratic processes and institutions’ (p. 1).

To conclude, the concept of observability emphasises the societal repercussions of platformisation and suggests public interest as a normative horizon for assessing and regulating them. It problematises the poor conditions for observing platform life and its effects, and suggests levelling off, in institutionalised ways, the information asymmetry between platforms and platform research. Thus, we think of observability as one possible ‘counter power’ in the sense of Helberger (2020, p. 9) who calls for establishing ‘entirely new forms of transparency’. First and foremost, observability therefore seeks to improve the informational conditions for studying the broader effects of platformisation. Over the next two sections, we discuss the modalities for such an approach.

3.2. Observe platform behaviour over time

Building on the arguments laid out in section two, the second principle of observability holds that the volatility of platforms requires *continuous* observation. While *ex ante* audits of technical mechanisms and *ex post* analysis of emblematic cases are certainly viable for more restricted systems, the dynamic and distributed nature of online platforms means that intermittent inspections or disclosures are insufficient, thwarted by the object’s transient character. Traditional forms of information sharing through transparency reports, legal inquiries, and regulated and structured disclosures, similar to those that exist for stock markets, can still be part of an observability framework, as can investigative reporting and whistleblowing. However, to tackle the specific challenges of digital platforms, more continuous forms of observation need to be envisaged.

When terms of service, technical design, or business practices change, the ‘rules of the game’ change as well, affecting platform participants in various ways. Projects like TOSBack⁷ use browser plugins and volunteer work to track and observe

7. <https://tosback.org>

changes in platforms' terms of service continuously, that is, while they are happening and not after some complaint has been filed. These are then distilled into more readable forms to accommodate wider audiences. The joint Polisis⁸ and PriBot⁹ projects pursue similar goals, drawing on artificial intelligence to interpret privacy policies and deal with the limitations of volunteer work. Such efforts should be made easier: a recent proposal by Cornelius (2019) suggests making terms of service contracts available as machine-readable documents to facilitate ongoing observation and interpretation. Similar approaches can be imagined for other areas of platform conduct, including technical tweaks or changes in business practices.

However, to account for the distributed and dynamic character of platform life, as it emerges from the interaction between policies, design choices, and use practices, continuous observation needs to reach beyond legal and technical specifications. Bringing the space of distributed outcomes into view is by no means easy, but the importance of doing so is increasingly clear. In their discussion of algorithms as policies, Hunt and McKelvey (2020, p. 330) indeed argue that the 'outcomes of these policies are as inscrutable as their intentions - under our current system of platform governance, it is beyond our reach to know whether algorithmic regulation is discriminatory or radicalizing or otherwise undermines the values that guide public policy'. Here, observability does not alter the underlying normative concerns but asks how platform reality can be sufficiently understood to make it amenable to normative reasoning in the first place. As platforms suck the bulk of online exchange into their increasingly centralised infrastructures, we need the capacity to probe not merely how algorithms work, but how fundamental social institutions are being reshaped. Answering these questions requires studying technical and legal mechanisms, use practices, and circulating units such as messages together. Given that our first goal is to understand rather than to place blame, there is no need to untangle networks of distributed causation from the outset. Entanglement and the wide variety of relevant questions we may want to ask mean that observability thus favours continuous and broad access to knowledge generating facilities.

There are at least four practical approaches that align with what we are aiming at. First, platforms have occasionally entered into *data access agreements* with researchers, journalists, NGOs, and so forth. Facebook is a case in point. The company's Data for Good¹⁰ programme, which builds 'privacy-preserving data products

8. <https://pribot.org/polisis>

9. <https://pribot.org>

10. <https://dataforgood.fb.com/>

to help solve some of the world's biggest problems', shares data with approved universities and civil society groups. The recently launched Social Science One initiative ¹¹, a collaboration with the US Social Science Council, is supposed to grant selected researchers access to both data and funding to study 'the impact of social media on elections and democracy' (King & Persily, 2019, p. 1). While these initiatives are good starting points, they have been plagued by delays and restrictions. Scholars have rightfully criticised that the scope and modalities for access remain in the hands of platforms themselves (Hegelich, 2020; Suzor et al., 2019). The central question is thus how to structure agreements in ways that asymmetries between platforms and third parties are reduced. Without a legal framework, companies can not only start and stop such initiatives at will but are also able to control parameters coming into play, such as thematic scope, coverage, and granularity.

Accountability interfaces providing continuous access to relevant data constitute a second direction. Facebook's Ad Library ¹², for example, is an attempt to introduce carefully designed observability, here with regard to (political) advertisement. Despite the limitations of the existing setup (see Leerssen et al., 2019), machine-readable data access for purposes of accountability can enable third-party actors to ask their own questions and develop independent analytical perspectives. While tools like Google Trends ¹³ are not designed for accountability purposes, a broader understanding of the term could well include tools that shed light on emergent outcomes in aggregate terms. There are already working examples in other domains, as the German Market Transparency Unit for Fuels ¹⁴, a division of the Federal Cartel Office shows. It requires gas stations to communicate current prices in real-time to make them available on the Web and via third-party Apps. ¹⁵ Well-designed data interfaces could both facilitate observability and alleviate some of the privacy problems other approaches have run into. One could even imagine sandbox-style execution environments that allow third parties to run limited code within platforms' server environment, allowing for privacy-sensitive analytics where data never leaves the server.

Developer APIs are data interfaces made available without explicit accountability purposes. These interfaces have been extensively repurposed to investigate the

11. <https://socialscience.one>

12. <https://www.facebook.com/ads/library/>

13. <https://trends.google.com>

14. https://www.bundeskartellamt.de/EN/Economicsectors/MineralOil/MTU-Fuels/mtufuels_node.html

15. <https://creativecommons.tankerkoenig.de/> / https://de.wikipedia.org/wiki/Markttransparenzstelle_für_Kraftstoffe

many social phenomena platforms host, ranging from political campaigning (e.g. Larsson, 2016) to crisis communication during disasters (e.g. Bruns & Burgess, 2014), as well as the technical mechanisms behind ranking and recommendation (e.g., Airolidi et al., 2016; Rieder et al., 2018). Depending on the platform, developer APIs provide data access through keyword searches, user samples, or other means. Twitter's random sample endpoint ¹⁶, which delivers representative selections of all tweets in real time (Morstatter et al., 2014), is particularly interesting since it allows observing overall trends while reducing computational requirements. One of the many examples for exploiting a data interface beyond social media is David Kriesel's project BahnMining ¹⁷, which uses the German railroad's timetable API to analyse train delays and challenge the official figures released by Deutsche Bahn.

But the so-called 'APIcalypse' (Bruns, 2019) that followed the Facebook-Cambridge Analytica scandal has led to restrictions in data access, rendering independent research much more difficult. Even before Facebook-Cambridge Analytica, working with developer APIs regularly created issues of reliability and reproducibility of results, research ethics, and privacy considerations (see Puschmann, 2019). Generally, developer interfaces are not designed for structured investigations into the layers of personalisation and localisation that may impact what users actually see on their screens. YouTube's 'up next' column is a case in point: while the API does make so-called 'related videos' available, it leaves out the personalized recommendations that constitute a second source for suggested videos. Research on the YouTube's recommender system, for example a study by PEW ¹⁸, is therefore necessarily incomplete. But the fact that developer APIs enable a wide variety of independent research on different topics means that in cases where privacy concerns can be mitigated, they are worth extending further. A structured conversation between platforms and research organisations about possible long-term arrangements is necessary and independent regulatory institutions could play a central role here.

Finally, due to API limitations, researchers have been relying on *scraping*, a set of techniques that glean data from end-user interfaces. Search engines, price snipers, and a whole industry of information aggregators and sellers rely on scraped data,

16. <https://developer.twitter.com/en/products/tweets/sample>

17. <https://www.heise.de/newsticker/meldung/36C3-BahnMining-offenbart-die-nackte-Wahrheit-hinter-der-DB-Puenktlichkeitsquote-4624384.html>

18. <https://www.pewinternet.org/2018/11/07/many-turn-to-youtube-for-childrens-content-news-how-to-lessons/>

but there are many non-commercial examples as well. Projects like AlgoTransparency¹⁹, run by former YouTube employee Guillaume Chaslot, regularly capture video recommendations from the web interface to trace what is being suggested to users. Roth et al. (2020) have recently used a similar approach to study whether YouTube indeed confines users to filter bubbles. Such high-profile questions call for empirical evidence, and since research results may change as quickly as systems evolve, continuous monitoring is crucial. While scraping does not demand active cooperation from the platforms under scrutiny, large-scale projects do require at least implicit acquiescence because websites can deploy a whole range of measures to thwart scraping.

Although more precarious than API-based approaches, taking data directly from the user interface allows for the explicit study of personalisation and localisation. Data retrieved through scraping may also serve to verify or critique data obtained through the previously mentioned techniques. Not unlike the panels assembled by analytics companies like Nielsen for their online products²⁰, the most promising platform-centred crowd-sourcing projects ask volunteers to install custom-built browser plugins to ‘look over their shoulder’. The *Datenspende* project, a collaboration between several German state-level media authorities, the NGO AlgorithmWatch, the Technical University Kaiserslautern, and Spiegel Online, recruited 4,500 volunteers before the German parliamentary elections in 2017 to investigate what users actually see when they look for party and candidate names on Google Search and Google News.²¹ The same approach was later used to scrutinise the SCHUFA²², Germany’s leading credit bureau, and most recently Instagram²³.

There are many other areas where scraping has been productively used. The \$heriff project²⁴, for example, also deployed browser plugins to investigate price discrimination practices on retail websites like Amazon (Iordanou et al., 2017). Even regulators have to resort to scraping: a recent study by the French *Conseil Supérieur de l’Audiovisuel* used the accounts of 39 employees and four fictitious users to study YouTube’s recommendation system.²⁵ The City of Amsterdam already began

19. <https://algotransparency.org>

20. <https://www.nielsen.com/us/en/solutions/measurement/online/>

21. <https://algorithmwatch.org/datenspende-unser-projekt-zur-bundestagswahl/>

22. <https://algorithmwatch.org/openschufa-warum-wir-diese-kampagne-machen/>

23. <https://algorithmwatch.org/instagram-algorithmus/>

24. <http://sheriff-v2.dynu.net/views/manual>

25. <https://www.csa.fr/Informer/Toutes-les-actualites/Actualites/Pourquoi-et-comment-le-CSA-a-realise-une-etude-sur-l-un-des-algorithmes-de-recommandations-de-YouTube>

scraping data from AirBnB in 2017²⁶, analysing consequences for the housing market and compliance by landlords with rules on short-term rentals. Given that sample quality, scale, and the dependence on platform acquiescence are significant disadvantages under current conditions, a legal framework regulating access to platform data would increase the practical viability of this approach. The current ambiguities risk creating chilling effects that discourage smaller research projects in particular. NYU's Ad Observer²⁷, a tool that uses browser plugins and scraping to investigate ad targeting on Facebook to compensate for the limitations of the above-mentioned Ad Library, tells a cautionary tale. The researchers recently received a cease and desist letter from the company, putting the whole project in peril (Horwitz, 2020).

However, it should be stated that not all forms of access to platform data further the public interest. Across all these four approaches we encounter serious privacy concerns. While there are areas where data access is unproblematic, others may require restricting access to certain groups, anonymise data, use aggregate statistics, or explore innovative models such as sandbox environments. These are not trivial problems; they raise the need for innovative and experimental approaches supported by institutional oversight. From a legal perspective, a recent interpretation of the GDPR by the European Data Protection Supervisor²⁸ clarified that research in the public interest must have leeway if done in accordance with ethical best practices. Still, concrete measures will need to be the subject of broader conversations about the appropriate balance to strike, which may lead, in certain cases, to more restrictions rather than fewer.

3.3. Strengthen capacities for collaborative knowledge creation

In his analysis of accountability as a social relation, Bovens (2007, p. 453) argues that 'transparency as such is not enough to qualify as a genuine form of accountability, because transparency does not necessarily involve scrutiny by a specific forum'. Given their deep and transversal impact, the question as to how knowledge about platforms is generated and how it circulates through society is crucial. In this section, we argue that effective accountability requires the participation of different actors and the generation of different forms of knowledge.

Our argument starts from the fact that platform companies have largely treated in-

26. <https://publicaties.rekenkamer.amsterdam.nl/handhaving-vakantieverhuurbestuurlijk-rapport/>

27. <https://adobserver.org>

28. https://edps.europa.eu/sites/edp/files/publication/20-01-06_opinion_research_en.pdf

formation about their systems, what users are posting or selling, and which kind of dynamics emerge from their interactions as private assets. They heavily invest in sophisticated analytics to provide insights and pathways for corporate action. Product development, optimisation, and detection and moderation of all kinds of illegal or ‘undesirable’ content have become important tasks that fully rely on evolving observational capabilities. While platforms would be able to facilitate knowledge creation beyond such operational concerns, the existing information asymmetries between those collecting and mining private data and society at large make this highly unlikely. Instead, platforms provide businesses and individual users with deliberately designed ‘market information regimes’ (Anand & Peterson, 2000) consisting of analytics products and services that provide information about the larger market and one’s own standing.

Creators on YouTube, for example, are now able to gauge how their videos are faring, how the choice of thumbnails affects viewer numbers, or how advertisers are bidding on keywords within the platform interface. But such interfaces are ‘socially and politically constructed and [...] hence fraught with biases and assumptions’ (Anand & Peterson, 2000, p. 270), privileging operational knowledge designed to boost performance over broader and more contextualised forms of insight. The narrow epistemological horizon of platform companies thus needs to be supplemented by inquiries that contextualise and question this business model. The problematic monopolisation of analytical capacities legitimises our demand for a more inclusive approach, which would open the locked-up data troves to qualified external actors. However, there simply is no one-size-fits-all approach able to cover all types of platforms, audiences, and concerns. Researchers, journalists, and activists are already engaged in ‘accountability work’, covering a range of questions and methods. Regulators add to this diversity: competition and antitrust inquiries require different forms of evidence than concerns regarding misinformation or radicalisation. We may therefore prefer to speak of ‘accountabilities’ in plural form.

There are many approaches coming from the technical disciplines that promise to enhance understanding. Emerging research fields like ‘explainable AI’ (e.g. Doran et al., 2017) seek to make primary ordering mechanisms more accountable, even if the issue remains of what ‘explainable’ means when different audiences ask different questions. Other strategies like the ‘glass box’ approach (Tubella & Dignum, 2019) focus on the monitoring of inputs and outputs to ‘evaluate the moral bounds’ of AI systems. A particularly rich example for image classification from Google Researchers comes in the form of an ‘activation atlas’, which intends to communicate how a convolutional neural network ‘sees’.²⁹ But since platforms are

much more than contained ordering mechanisms, the problem of how to make their complexity readable, how to *narrate* what can be gleaned from data (see Dourish, 2016), remains unsolved. However, researchers in the humanities and social sciences have long been interested in how to make sense of quantitative information. Work on ‘narrating numbers’ (Espeland, 2015), ‘narrating networks’ (Bounegru et al., 2017), or the substantial research on information visualisation (e.g. Drucker, 2014) can serve as models. But as Sloane & Moss (2019) argue in their critique of current approaches to AI, there is a broader ‘social science deficit’ and the one-sided focus on quantitative information is part of the problem. The marginalisation of qualitative methods such as ethnographic work that tries to elucidate both the context within which platforms make decisions and the meaning actors ascribe to practices and their effects, limits knowledge production.

Journalists also have unique expertise when it comes to forms of knowledge generation and presentation. A recent example is the work by Karen Hao and Jonathan Stray³⁰ on the controversial KOMPASS project,³¹ which questions the very possibility of fair judgements by allowing users to ‘play’ with the parameters of a simplified model. Likewise, NGOs have long worked on compound forms of narration that combine different data sources and methods for purposes of accountability. Greenpeace’s Guide to Greener Electronics, which includes a grade for companies’ willingness to share information, or the Ranking Digital Rights³² project are good examples for the translation of research into concrete political devices. Accountability, understood as an inherent element of democratic control, cannot be reduced to a forensic process that transposes ‘facts’ from obscurity into the light. It needs to be considered as an ongoing social achievement that requires different forms of sense-making, asking for contributions from different directions and epistemological sensitivities. Access to machine-readable data, our focus in the last section, has limitations, but also allows different actors to develop their own observation capacities, adapting their analytical methods to the questions they want to ask.

We are aware that increased understanding of platform life would prompt reactions and adaptations by different stakeholders gathering around platforms, including actors seeking to ‘game’ the system and even platform owners themselves.

29. <https://distill.pub/2019/activation-atlas/>

30. <https://www.technologyreview.com/s/613508/ai-fairer-than-judge-criminal-risk-assessment-algorithm/>

31. <https://www.technologyreview.com/s/607955/inspecting-algorithms-for-bias/>

32. <http://rankingdigitalrights.org>

Making the constant negotiations between these actors more visible may have the advantage, however, that the process of establishing boundaries of acceptable behaviour could be engaged more explicitly. As Ziewitz (2019, p. 713) argues for the field of search engine optimisation (SEO), ‘the moral status of reactive practices is not given, but needs to be accomplished in practice’. Distributing this ‘ethical work’ over a wider array of actors could thus be a step toward some modest form of ‘co-operative responsibility’ (Helberger et al., 2018), even if fundamental power asymmetries remain.

Observability thus raises the complicated question of how data and analytical capacities should be made available, to whom, and for what purpose. This clearly goes beyond data access. As Kemper & Kolkman (2019) note, ‘no algorithmic accountability without a critical audience’, and the capacity for critique requires more than a critical attitude. For this reason, frameworks for data access should ‘go hand-in-hand with the broader cultivation of a robust and democratic civil society, which is adequately funded and guaranteed of its independence’ (Ausloos et al., 2020, p. 86). And Flyverbom (2015, p. 115) reminds us that transparency, understood as a transformative process, cannot succeed ‘without careful attention to the formats, processes of socialization, and other affordances of the technologies and environments in which they play out’. Monitoring platforms on a continuous basis may thus call for considerable resources if done well. Governmental institutions, possibly on a European level, could play a central role in managing data access, in making long-term funding available for research, and in coordinating the exchange between existing initiatives. But given the complexity of the task, regulators will also have to build ‘in-house’ expertise and observational capacity, backed by strong institutional support.

The capacity to make sense of large and complex socio-technical systems indeed relies on a number of material conditions, including access to data, technical expertise, computing power, and not least the capacity to connect data-analytical practices to social concerns. Such a capacity is typically produced as a collective effort, through public discourse. The quality of observability depends on such discourses to explore what kind of knowledge forms allow concerned actors to make actually meaningful interpretations.

4. Conclusion: toward platform observability

This article developed the concept of observability to problematise the assumptions and expectations that drive our demands for transparency of platform life. Observability is not meant to be a radical departure from the call for transparency.

Rather, it draws practical conclusions from the discrepancy we noted between the complexity of the platform machinery and the traditional idea of shedding light on and seeing as a way of establishing external oversight. In a nutshell, we are suggesting observability as a pragmatic, knowledge-focused approach to accountability. Observability stresses technical and social complexities, including the distributed nature of platform behaviour. Moreover, it regards continuous and collaborative observation within a normative framework as a necessary condition for regulating the explosive growth of platform power. We see three main directions where further steps are needed to move closer to the practical realisation of these principles.

Regulating for observability means working toward structured information interfaces between platforms and society.³³ To account for quickly changing circumstances, these interfaces need to enable continuous observation. To allow for a broader set of questions to be asked, a broad range of data has to be covered. And to bring a wider variety of epistemological sensitivities into the fold, they need to be sufficiently flexible. What constitutes suitable and sufficient access will have to be decided on a per-platform basis, including the question of who should be able to have access in the first place. But the examples we briefly discussed in section 3.2—and the many others we left out—show that there is already much to build on. The main goal, here, is to develop existing approaches further and to make them more stable, transparent, and predictable. Twitter's new API³⁴, which now explicitly singles out academic research use cases, is a good example for a step in the right direction, but these efforts are still voluntary and can be revoked at any time. Without binding legal frameworks, platforms can not only terminate such initiatives at will, they also control relevant modalities such as thematic scope and depth of access. Realigning the structural information asymmetries between platforms and society thus requires curtailing the de facto ownership over data that platforms collect about their users.

Observability as part of regulation requires engaging with the specific properties of algorithmic systems and the co-produced nature of platform behaviour. The complex interactions between technical design, terms of service, and sometimes vast numbers of both users and 'items' mean that the concept of a singular algorithm steering the ordering processes at work in large-scale platforms is practically and conceptually insufficient. If techniques like machine learning are here to stay, regulatory approaches will have to adapt to conditions where the object of regulation

33. This aligns with Sandvig et al. (2014, p. 17), who call for 'regulation toward auditability'.

34. https://blog.twitter.com/developer/en_us/topics/tools/2020/introducing_new_twitter_api.html

is spread out, volatile, and elusive. The pressing questions are not restricted to how and what to regulate, but also encompass the issue of what platforms are doing in the first place. While normative concepts such as algorithmic fairness or diversity are laudable goals, their focus seems rather narrow considering the fundamental change of markets and the public sphere that platforms provoke. We therefore suggest the broader concept of public interest as a normative benchmark for assessing platform behaviour, a concept obviously in need of specification. But whatever set of norms or values are chosen as guiding principles, the question remains how to ‘apply’ them, that is, how to assess platform behaviour against public interest norms. Observation as a companion to regulation stresses the fact that we need to invest in our analytical capacities to undergird the regulatory response to the challenges platforms pose. Likewise, the existing approaches to studying platforms should be supplemented with specific rights to information. Together, these elements would constitute important steps towards a shared governance model (see Helberger et al., 2018), where power is distributed more equally between platforms and their constituencies.

Institutionalising processes of collective learning refers to the need to develop and maintain the skills that are required to observe platforms. A common characteristic of the data collecting projects mentioned above is their ephemeral, experimental, and somewhat amateurish nature. While this may sound harsh, it should be obvious that holding platforms to account requires ‘institution-building’, that is, the painstaking assembly of skills and competence in a form that transposes local experiments into more robust practices able to guarantee continuity and accumulation. While academic research fields have their own ways of assembling and preserving knowledge, the task of observing large-scale platforms implies highly specialised technical and logistical feats that few organisations are able to tackle. Material resources are only one part of the equation and the means to combat discontinuity and fragmentation are at least equally important. One form of institutional incorporation of observability would therefore be something akin to ‘centres of expertise’ tasked with building the capacity to produce relevant knowledge about platforms. Such centres could act as an, ‘important bridge builder between those holding the data and those wishing to get access to that data’ (Ausloos et al., 2020, p. 83). Pushing further, a European Platform Observatory,³⁵ driven by a public interest mandate, equipped with adequate funding, and backed by strong regu-

35. The European Commission is already hosting an Observatory on the Online Platform Economy (<https://platformobservatory.eu/>) – of which both authors are members – and it plans to create a digital media observatory. <https://ec.europa.eu/digital-single-market/en/news/commission-launches-call-create-european-digital-media-observatory>. However, both bodies have a thematically restricted mandate and lack any regulatory authority.

latory support, could be a way forward to platform accountability.

Holding platforms to account is a complex task that faces many challenges. However, given their rising power, it is quickly becoming a necessity. The concept of observability spells out these challenges and suggests steps to tackle them, taking a pragmatic, knowledge-based approach. The goal, ultimately, is to establish observability as a ‘counter power’ to platforms’ outsized hold on contemporary societies.

ACKNOWLEDGEMENTS

This work was, in part, inspired by discussions we had as members of the European Commission’s Observatory on the Online Platform Economy. We would also like to thank Joris van Hoboken, Paddy Leerssen, and Thomas Poell for helpful comments and feedback.

References

- Agre, P. E. (1994). Surveillance and Capture: Two Models of Privacy. *The Information Society*, 10(2), 101–127. <https://doi.org/10.1080/01972243.1994.9960162>
- Albu, O. B., & Flyverbom, M. (2019). Organizational Transparency: Conceptualizations, Conditions, and Consequences. *Business & Society*, 58(2), 268–297. <https://doi.org/10.1177/0007650316659851>
- Anand, N., & Peterson, R. A. (2000). When Market Information Constitutes Fields: Sensemaking of Markets in the Commercial Music Industry. *Organization Science*, 11(3), 270–284. <https://doi.org/10.1287/orsc.11.3.270.12502>
- Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, 20(3), 973–989. <https://doi.org/10.1177/1461444816676645>
- August, V., & Osrecki, F. (2019). Transparency Imperatives: Results and Frontiers of Social Science Research. In V. August & F. Osrecki (Eds.), *Der Transparenz-Imperativ: Normen – Praktiken – Strukturen* (pp. 1–34). Springer. <https://doi.org/10.1007/978-3-658-22294-9>
- Bernstein, E. S. (2012). The Transparency Paradox: A Role for Privacy in Organizational Learning and Operational Control. *Administrative Science Quarterly*, 57(2), 181–216. <https://doi.org/10.1177/0001839212453028>
- Bogost, I. (2015, January 15). The Cathedral of Computation. *The Atlantic*. <https://www.theatlantic.com/technology/archive/2015/01/the-cathedral-of-computation/384300/>
- Bovens, M. (2007). Analysing and Assessing Accountability: A Conceptual Framework. *European Law Journal*, 13(4), 447–468. <https://doi.org/10.1111/j.1468-0386.2007.00378.x>
- Brandeis, L. D. (1913, December 20). What publicity can do. *Harper’s Weekly*.

- Bruns, A. (2019). After the 'APIcalypse': Social media platforms and their fight against critical scholarly research. *Information, Communication & Society*, 22(11), 1544–1566. <https://doi.org/10.1080/1369118X.2019.1637447>
- Bruns, A., & Burgess, J. (2013). Crisis communication in natural disasters: The Queensland floods and Christchurch earthquakes. In K. Weller, A. Bruns, J. Burgess, M. Mahrt, & C. Puschmann (Eds.), *Twitter and Society* (pp. 373–384). Peter Lang.
- Burrell, J. (2016). How the machine "thinks": Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 1–12. <https://doi.org/10.1177/2053951715622512>
- Cohen, J. E. (2019). *Between Truth and Power: The Legal Constructions of Informational Capitalism*. Oxford University Press. <https://doi.org/10.1093/oso/9780190246693.001.0001>
- Cornelius, K. B. (2019). Zombie contracts, dark patterns of design, and 'documentisation'. *Internet Policy Review*, 8(2). <https://doi.org/10.14763/2019.2.1412>
- Covington, P., Adams, J., & Sargin, E. (2016). Deep Neural Networks for YouTube Recommendations. *Proceedings of the 10th ACM Conference on Recommender Systems*, 191–198. <https://doi.org/10.1145/2959100.2959190>
- Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, 55(10), 78–87. <https://doi.org/10.1145/2347736.2347755>
- Doran, D., Schulz, S., & Besold, T. R. (2017). What Does Explainable AI Really Mean? A New Conceptualization of Perspectives. *ArXiv*. <http://arxiv.org/abs/1710.00794>
- Douglass, B. (1980). The Common Good and the Public Interest. *Political Theory*, 8(1), 103–117. <http://doi.org/10.1177/009059178000800108>
- Dourish, P. (2016). Algorithms and their others: Algorithmic culture in context. *Big Data & Society*, 3(2). <https://doi.org/10.1177/2053951716665128>
- Espeland, W. (2015). Narrating Numbers. In R. Rottenburg, S. E. Merry, S.-J. Park, & J. Mugler (Eds.), *The World of Indicators: The Making of Governmental Knowledge through Quantification* (pp. 56–75). Cambridge University Press. <https://doi.org/10.1017/CBO9781316091265.003>
- Etzioni, A. (2010). Is Transparency the Best Disinfectant? *Journal of Political Philosophy*, 18(4), 389–404. <https://doi.org/10.1111/j.1467-9760.2010.00366.x>
- Ezrahi, Y. (1992). Technology and the civil epistemology of democracy. *Inquiry*, 35(3–4), 363–376. <https://doi.org/10.1080/00201749208602299>
- Flyverbom, M. (2016). Transparency: Mediation and the Management of Visibilities. *International Journal of Communication*, 10, 110–122. <https://ijoc.org/index.php/ijoc/article/view/4490>
- Fourcade, M., & Gordon, J. (2020). Learning Like a State: Statecraft in the Digital Age. *Journal of Law and Political Economy*, 1(1), 78–108. <https://escholarship.org/uc/item/3k16c24g>
- Gillespie, T. (2018). *Custodians of the Internet*. Yale University Press.
- Hegelich, S. (2020). Facebook needs to share more with researchers. *Nature*, 579, 473–473. <https://doi.org/10.1038/d41586-020-00828-5>
- Helberger, N. (2020). The Political Power of Platforms: How Current Attempts to Regulate Misinformation Amplify Opinion Power. *Digital Journalism*, 8(3). <https://doi.org/10.1080/21670811.2020.1773888>

- Helberger, N., Pierson, J., & Poell, T. (2018). Governing online platforms: From contested to cooperative responsibility. *The Information Society*, 34(1), 1–14. <https://doi.org/10.1080/01972243.2017.1391913>
- Horwitz, J. (2020, October 23). Facebook Seeks Shutdown of NYU Research Project Into Political Ad Targeting. *The Wall Street Journal*. <https://www.wsj.com/articles/facebook-seeks-shutdown-of-nyu-research-project-into-political-ad-targeting-11603488533>
- Hunt, R., & McKelvey, F. (2019). Algorithmic Regulation in Media and Cultural Policy: A Framework to Evaluate Barriers to Accountability. *Journal of Information Policy*, 9, 307–335. <https://doi.org/10.5325/jinfopoli.9.2019.0307>
- Iordanou, C., Soriente, C., Sirivianos, M., & Laoutaris, N. (2017). Who is Fiddling with Prices?: Building and Deploying a Watchdog Service for E-commerce. *Proceedings of the Conference of the ACM Special Interest Group on Data Communication - SIGCOMM*, 17, 376–389. <https://doi.org/10.1145/3098822.3098850>
- Kemper, J., & Kolkman, D. (2019). Transparent to whom? No algorithmic accountability without a critical audience. *Information, Communication & Society*, 22(14), 2081–2096. <https://doi.org/10.1080/1369118X.2018.1477967>
- King, G., & Persily, N. (2019). A New Model for Industry–Academic Partnerships. *PS: Political Science & Politics*, 53(4), 703–709. <https://doi.org/10.1017/S1049096519001021>
- Langley, P., & Leyshon, A. (2017). Platform capitalism: The intermediation and capitalisation of digital economic circulation. *Finance and Society*, 3(1), 11–31. <https://doi.org/10.2218/fin soc.v3i1.1936>
- Larsson, A. O. (2016). Online, all the time? A quantitative assessment of the permanent campaign on Facebook. *New Media & Society*, 18(2), 274–292. <https://doi.org/10.1177/1461444814538798>
- Leerssen, P., Ausloos, J., Zarouali, B., Helberger, N., & Vreese, C. H. (2019). Platform Ad Archives: Promises and Pitfalls. *Internet Policy Review*, 8(4), 1–21. <https://doi.org/10.14763/2019.4.1421>
- Lessig, L. (1999). *Code: And other laws of cyberspace*. Basic Books.
- Mattli, W., & Woods, N. (2009). In Whose Benefit? Explaining Regulatory Change in Global Politics. In W. Mattli & N. Woods (Eds.), *The Politics of Global Regulation* (pp. 1–43). <https://doi.org/10.1515/9781400830732.1>
- Morgan, B., & Yeung, K. (2007). *An introduction to Law and Regulation*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511801112>
- Morstatter, F., Pfeffer, J., & Liu, H. (2014). When is it Biased? Assessing the Representativeness of Twitter's Streaming API. *ArXiv*. <http://arxiv.org/abs/1401.7909>
- Napoli, P. M. (2015). Social media and the public interest: Governance of news platforms in the realm of individual and algorithmic gatekeepers. *Telecommunications Policy*, 39(9), 751–760. <https://doi.org/10.1016/j.telpol.2014.12.003>
- Obar, J. A. (2020). Sunlight alone is not a disinfectant: Consent and the futility of opening Big Data black boxes (without assistance). *Big Data & Society*, 7(1). <https://doi.org/10.1177/2053951720935615>
- Parsons, C. (2015). Beyond Privacy: Articulating the Broader Harms of Pervasive Mass Surveillance. *Media and Communication*, 3(3), 1–11. <https://doi.org/10.17645/mac.v3i3.263>

Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press.

Power, M. (1997). *The audit society. Rituals of verification*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198296034.001.0001>

Powles, J., & Nissenbaum, H. (2018, December 7). The Seductive Diversion of ‘Solving’ Bias in Artificial Intelligence. *OneZero*. <https://onezero.medium.com/the-seductive-diversion-of-solving-bias-in-artificial-intelligence-890df5e5ef53>

Puschmann, C. (2019). An end to the wild west of social media research: A response to Axel Bruns. *Information, Communication & Society*, 22(11), 1582–1589. <https://doi.org/10.1080/1369118X.2019.1646300>

Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J.-F., Breazeal, C., Crandall, J. W., Christakis, N. A., Couzin, I. D., Jackson, M. O., Jennings, N. R., Kamar, E., Kloumann, I. M., Larochelle, H., Lazer, D., McElreath, R., Mislove, A., Parkes, D. C., Pentland, A. ‘Sandy’, ... Wellman, M. (2019). Machine behaviour. *Nature*, 568(7753), 477–486. <https://doi.org/10.1038/s41586-019-1138-y>

Rieder, B. (2020). *Engines of Order. A Mechanology of Algorithmic Techniques*. Amsterdam University Press. <https://doi.org/10.2307/j.ctv12sdvf1>

Rieder, B., Matamoros-Fernández, A., & Coromina, Ò. (2018). From ranking algorithms to ‘ranking cultures’: Investigating the modulation of visibility in YouTube search results. *Convergence*, 24(1), 50–68. <https://doi.org/10.1177/1354856517736982>

Roth, C., Mazières, A., & Menezes, T. (2020). Tubes and bubbles topological confinement of YouTube recommendations. *PLOS ONE*, 15(4). <https://doi.org/10.1371/journal.pone.0231703>

Rowland, N. J., & Passoth, J.-H. (2015). Infrastructure and the state in science and technology studies. *Social Studies of Science*, 45(1), 137–145. <https://doi.org/10.1177/0306312714537566>

Sandvig, C., Hamilton, K., Karahalios, K., & Langbort, C. (2014). *Auditing algorithms: Research methods for detecting discrimination on internet platforms*. Data and Discrimination: Converting Critical concerns into productive inquiry, a Preconference at the 64th Annual Meeting of the International Communication Association, Seattle, WA. <https://pdfs.semanticscholar.org/b722/7cbd34766655dea10d0437ab10df3a127396.pdf>

Seaver, N. (2017). Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data & Society*, 4(2), 1–12. <https://doi.org/10.1177/2053951717738104>

Sloane, M., & Moss, E. (2019). AI’s social sciences deficit. *Nature Machine Intelligence*, 1(8), 330–331. <https://doi.org/10.1038/s42256-019-0084-6>

Staab, P. (2019). *Digitaler Kapitalismus: Markt und Herrschaft in der Ökonomie der Unknappheit*. Suhrkamp.

Stolton, S. (2019, November 20). Vestager takes aim at ‘biopower’ of tech giants. *EURACTIV*. <https://www.euractiv.com/section/copyright/news/vestager-takes-aim-at-biopower-of-tech-giants/>

Suchman, L. A. (2007). *Human-Machine Reconfigurations. Plans and Situated Actions* (Second). Cambridge University Press. <https://doi.org/10.1017/CBO9780511808418>

Suzor, N. P., Myers West, S., Quodling, A., & York, J. (2019). What Do We Mean When We Talk About Transparency? Toward Meaningful Transparency in Commercial Content Moderation. *International Journal of Communication*, 13, 1526–1543. <https://ijoc.org/index.php/ijoc/article/view/9736/0>

van Dijck, J. (2020). Governing digital societies: Private platforms, public values. *Computer Law & Security Review*, 36. <https://doi.org/10.1016/j.clsr.2019.105377>

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

Yeung, K. (2017). 'Hypernudge': Big Data as a mode of regulation by design. *Information, Communication & Society*, 20(1), 118–136. <https://doi.org/10.1080/1369118X.2016.1186713>

Ziewitz, M. (2019). Rethinking gaming: The ethical work of optimization in web search engines. *Social Studies of Science*, 49(5), 707–731. <https://doi.org/10.1177/0306312719865607>

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

Published by



ALEXANDER VON HUMBOLDT
INSTITUTE FOR INTERNET
AND SOCIETY

in cooperation with



CREATE

centre
— internet
et —
societe



R&I
IN3
Internet
interdisciplinary
Institute
Universitat Oberta de Catalunya