This is a postprint version of the article published as:
Shin, D., Hameleers, M., Park, Y.Y., Jeong, N.K., Triell, D., Diakopolous, N., Helberger, N., Lewis, S.C., Westlund, O. & Baumann, S. (2022). Countering Algorithmic Bias and Disinformation and Effectively Harnessing the Power of AI in Media. Journalism & Mass Communication Quarterly. 2022, 99 (4), 887-907

Artificial intelligence (AI) is affecting the daily lives of billions of media users (Wölker & Powell, 2021). Algorithms are popular and effective tools utilized by companies online, but their popularity comes at the expense of systematic discrimination, limited transparency, and vague accountability (Moller et al., 2018). Algorithmic filtering procedures may lead to more impartial, and thus possibly fairer processes than those processed by humans. However, algorithmic recommendation processes have been criticized for their tendency to intensify/reproduce bias, distortion of facts, information asymmetry, and process opacity (Ananny & Crawford, 2018). Algorithmic bias may deteriorate algorithmic injustice that machine learning automates and perpetuates unjust and discriminatory patterns (Hoffman, 2019).

Recent algorithmic platforms have faced similar dilemmas (Shin, 2021). While algorithmic platforms offer personalized and relevant content in innovative interactive ways, the ethical and privacy issues are complicated and intertwined with algorithmic personalization (Helberger, Karppinen, & D'Acunto, 2018). Questions regarding how to safeguard the goals, values, and personalizing processes of algorithms, to what extent users need to share personal information with algorithms, and how to balance privacy and algorithmic personalization remain controversial (Crain, 2018). Underlying these questions are concerns about how to mitigate bias and discrimination in data, as well as urgent tasks on how to design algorithmic platforms that are transparent and fair (Hoffman et al., 2019). As ethical concerns have peaked recently with the rise of algorithmic media, the opacity of black-box algorithm processes had led to calls for studies on fairness and transparency (Dörr & Hollnbuchner, 2017).

JMCQ Invited Forum

Recent research (e.g., Park, 2019; Sandvig et al., 2016) has highlighted normative implications and problems associate with these algorithms when it comes to *fairness*, *accountability*, and *transparency*. Transparency and fairness particularly emerge as key attributes of trustworthy algorithmic systems processing user-sensitive data (Helberger et al., 2018). This topic will be even more critical when media platforms utilize more and more sophisticated algorithms and people rely more on algorithms than social influence when making judgments. AI is becoming pervasive across all media industries and service functions. This transformation brings to the fore several key questions: How to govern these algorithms effectively and legitimately while ensuring that they are user-centered and socially responsible? How can users make sense of algorithmic fairness and how do they construe algorithmic transparency? How do users perceive algorithm-based processes in general? (Graefe et al., 2018). As these normative concerns have given rise to calls for a better explanatory framework to address them (Thurman et al., 2019), a number of studies have examined these concerns from various perspectives, such as a user consumption perspective (how people make sense of it in their everyday lives), ethical point of view (how journalism practices face and deal with the ethical issues), and regulatory and managerial perspective (how to govern and manage algorithmic bias effectively).

The purpose of this Invited Forum is to continue this discussion by inviting leading scholars in the area to share their views on algorithmic biases. Our forum is set to contribute to theorizing and operationalizing algorithmic media platforms that are fairer, more transparent, and more responsible. To this end, this forum aims to contribute to the understanding of algorithmic bias, leading to operational, user-centric definitions for different areas of media platforms with implications for both design/developments and sociological/ethical models. The invited essays in this forum offer theoretical insights into user information processing through clarification of algorithmic sensemaking processing. The forum highlights the role of transparent fairness as part of broader considerations of ethics by design in algorithmic media.

Michael Hameleers, Assistant Professor at the University of Amsterdam, starts off our forum discussion and responds to the call to research algorithmic amplification. He exemplifies algorithmic amplifications with Covid-19 vaccines that citizens search for information on the vaccines motivated by doubts on its effectiveness, their search behavior and prior exposure patterns can cultivate a restricted media environment where a disproportionate amount of conspiracies, deceptive content, and clickbait reinforcing their doubt is offered – amplifying people's existing doubts and motivating the selection of increasingly less diverse, attitude-reinforcing content. He proposes three directions for future research of algorithmic biases in media effect studies: (1) move beyond voluntary selective exposure and confirmation biases; (2) take a holistic approach to effect studies; and (3) look at the longer-term and spiraling impact of algorithmic amplification. He concludes with urgent calls to design algorithmic media that are transparent and fair that are thus crucial in the fight against disinformation but should be supplemented with educational packages and tools that offer users more control.

Yong Jin Park, Professor at Howard University, and Jeong Nam Kim, Professor at the University of Oklahoma, rightly point out the dual aspect of algorithmic personalization. Personalization through AI inherently involves the consideration of privacy and amplification of built-in bias in their recommended contents. They argue that algorithmic personalization is a construct of deliberate choices as algorithmic platforms set up the conditions with finite choices that proximate user preferences. A profound myth to debunk is that it is possible to detect *what users want* and deliver *what they need*, as if there were the intrinsic relationship between the accuracy of the personalization and the extent of data collection – the false premise of the validity of algorithmic personalization. Finally, Park and Kim conclude that the algorithmic bias debate should start with the realistic acknowledgment that purely mathematical solutions will never reach the optimal point of fairness. It has been highlighted realistic assessment regarding algorithmic (in)capability in its uninvited custodian roles, which cannot resolve complex ethical objectives by having more data or statistical model choices. This algorithmic incapability is a nice segway into discussing algorithms from a socio-technical perspective.

Nick Diakopoulos and Daniel Trielli at Northwestern University approach algorithmic media as a complex socio-technical system that includes different actors whose motivations, pressures, and values influence the curation of content. They describe two lenses to analyze the interplay between actors in algorithmic media systems: power and values. These lenses enable us to look at a meaningful interpretation of curation in algorithmic media by examining not only at the algorithm or the users in isolation, but also at the intersections in curation processes between algorithmic mediators, the actions of the users, and the availability of content provided by publishers or creators. These arguments are very persuasive because simply describing the actors in algorithmic media curation make it hard to untangle one from the other: the algorithm is shaped by the values imbued in it by designers, but also by the needs of the user and the opportunities of the source media; the user shapes and responds to the algorithm in search for the available content; the available content is shaped by the needs of the user as refracted through the pressures of the algorithm as understood by the creators. Their key arguments resonate with Helberger, Lewis, and Westlund's views.

Natali Helberger, Professor at the University of Amsterdam, Seth Lewis, Professor at the University of Oregon, and Oscar Westlund, Professor at Oslo Metropolitan University discuss the responsibility of AI from a broader socio-technical system by conceptualizing responsibility as connected to moral and legal requirements associated with any agent in the system. Building on the four A's socio-technical framework, they approach AI as a socio-technical system that, as a key technological *actant*, intersects with social *actors* that are central to the shaping of media production as well as diverse *audiences*. These actors, actants, and audiences are interconnected through the *activities* of media work, and it's in the interactions that occur among them—the stuff of which a socio-technical system is made.

Sabine Baumann, Professor at Jade University of Applied Sciences in Germany, challenges existing pop-culture notions of AI technologies and investigates the human element of supposed AI failures. She also notes the black box phenomenon of AI and calls for research on explainable AI so that AI enables users to appropriately understand and effectively manage the AI systems as a prerequisite of a trustworthy AI. While Baumann proposes numerous remedial strategies for AI to be responsible and fair, she predicts that super-intelligent general AI systems will not happen any time soon. This resonates with Trielli and Diakopoulos' argument that AI cannot fully replace humans because AI judgment relies purely on trained logic while humans use empathy, imagination, and valuation. This view is in line with Hameleers who emphasized the importance of educating news users how they can control algorithms, and how their own choices may afford them more control over the techniques concerning their choices. This can be a realm of algorithmic literacy (Shin et al., 2021). All authors seem to unanimously suggest that human moral responsibility and proactive user control as one of the key factors avoiding bias in AI technologies.

Don Donghee Shin, Ph.D.

Invited Forum Editor

The Algorithmic Amplification of Dishonesty and Biased Realities: Challenges, Implications and A Future Research Agenda

Disinformation has been regarded as a key threat to democracies throughout the globe (e.g., Bennett & Livingston, 2017; Marwick & Lewis, 2017). The phenomenal spread of deceptive information cannot be understood without taking the context of digital affordances into account (e.g., Starbird et al., 2019; Zhang et al., 2021). Online, citizens can easily find a version of reality that best fits their identities and beliefs. This process, however, may not fully operate in a transparent, controlled, and voluntary way: Disinformation may find susceptible recipients and augment their beliefs through algorithmic amplification. This means that existing doubts and selection patterns limit people's selection options and create an illusionary worldview of like-mindedness – legitimizing deceptive narratives, conspiracies and untruths. In that sense, the algorithmic recommendation of deceptive content may impede learning from the other side, create an illusion of social support, and herewith reproduces (dis)information biases (Ananny & Crawford, 2018).

Why would this be a challenge to democracy, as algorithms and recommendation systems should offer tailor-made recommendations on journalistic products that best fit people's interests and beliefs? The problem is that algorithms create an uneven playing field without informing news users on its workings, biases, and implications (Helberger et al., 2018; Möller et al., 2018). To offer an example, when citizens search for information on COVID-19 vaccines motivated by doubts on its effectiveness, their search behavior and prior exposure patterns can cultivate a restricted media environment where a disproportionate amount of conspiracies, deceptive content, and clickbait reinforcing their doubt is offered – amplifying people's existing doubts and motivating the selection of increasingly less diverse, attitude-reinforcing content.

As argued by Waisbord (2018), eroding levels of trust and the politicization of science may have contributed to an ‘epistemic democracy’ where multiple contesting truth claims compete for attention and legitimacy. In this setting, existing (moderate) levels of uncertainty and distrust may motivate the selection of attitude-reinforcing disinformation. The processing of such information further strengthens uncertainty and distrust. This feedback-loop is amplified by algorithms in the digital media environment, which promote attitude-consistent information selection and limit cross-cutting news options. Although algorithms on their own may not be as dangerous and autonomous as assumed in some dystopian views, they may amplify existing fears, distrust, and confirmation biases.

These shifts in the digital ecology of factual relativism create a breeding ground for the rejection of the scientific paradigm, conspiracy theories, and attacks on scientific consensus (Waisbord, 2018). This is highly problematic as the epistemic foundations of deliberative democracy are at stake (Arendt, 1967). Hence, reasoned from people’s own biased news ecologies, the illusion of truth and social support is amplified because algorithmic amplification has created – without acknowledging its biases – a worldview governed by false consensus, social support, and evidence, herewith transgressing the model of a well-functioning deliberative democracy where citizens can learn from the other side and engage in argumentative debates. What’s left to debate if everyone seems to agree with you?

What Should (We/They) Do About It?

Although notions of omnipresent ‘filter bubbles’ and ‘echo chambers’ have been debated for quite some time (see e.g., Zuiderveen Borgesius et al., 2016), we can and should not ignore the fact that algorithms do, at least for vulnerable segments of the news audience, reinforce spirals of deceptive content and niches of disinformed audiences clinging on to alternative worldviews. The specific problem with false information is that lying and deceiving is a fundamental right that different actors may exercise. Legal boundaries are only crossed when deception takes on the shape of hate speech or discriminatory language. A simple ban on disinformation is thus not feasible, nor desirable. In my view, a more viable and sustainable solution is to reveal the mechanisms of algorithms, offering people insights into the workings of the ‘black box’ and educating them on their implications. In addition, user control should be built into the dynamics of recommendation systems – allowing news users the choice if, and if so, to what extent, they would like their exposure options to be tailored by recommender systems.

Regulators, policymakers, and platforms should all make important steps in the direction of transparency, accountability, and education. Urgent calls to design algorithmic systems that are transparent and fair (Hoffman et al., 2019) are thus crucial in the fight against disinformation but should be supplemented with educational packages and tools that offer users more control. It is not enough to simply reveal what’s hidden in the black box without educating news users on how they can act on this, and how their own choices may afford them more control over the techniques concerning their choices.

A Research Agenda Integrating Algorithmic Amplification in Media Effect Studies

To arrive at these solutions, we need recommendations that are founded on empirical evidence on the workings (and potentially discriminatory) implications of algorithms. I, therefore, propose three directions for future research integrating algorithmic biases in media effect studies:

1. Move beyond ‘voluntary’ selective exposure and confirmation biases. Most research looking at selective exposure and avoidance assumes that biased information environments are more or less driven by individuals’ own choices and automatic processing biases. Yet, algorithms create an uneven playing field by limiting the number of options people can select or avoid, and these biasing and discriminatory practices should be taken into account to fully comprehend how people select (depictive) content in fragmented information settings.

2. Take a holistic approach to effect studies. We should not blame it all on the algorithm (Möller et al., 2018). People's biased news diets may start with their own doubts, uncertainty, and fear – and the neglect of their sentiments by the established press. Hence, disinformation's effectiveness is for a large part driven by distrust in mainstream media and the feeling of being unrepresented (e.g., Zimmermann & Kohring, 2020). Although such feelings may be amplified by algorithms, they have a socio-political origin that should not be neglected.

3. Look at the longer-term and spiraling impact of algorithmic amplification. The real democratic and societal impact of algorithms and their amplifying nature does not exist after one exposure moment. It rather consists of a reinforcing spiral that – over time – places people in niches or traps of deceptive content that will increasingly resonate with their own beliefs. For this reason, future research should map the impact of algorithmic amplification over time, assessing how people's information ecology gradually becomes less diverse, more consistent, and more extreme.

Conclusion

Although algorithms may be conducive to a media ecology that is more personalized and tailored to individual needs and preferences, they may also threaten democratic communication by reinforcing disinformed worldviews and likeminded information. Here, I argue that algorithms are not the cause of disinformation or polarization, but rather a catalyst or amplifying factor: Existing doubts and distrust may be strengthened when algorithms recommend like-minded content whilst filtering out discrepant or nuanced views that could help them to relativize their distrust and counter misperceptions. Importantly, future research needs to map how algorithms play a role in limiting people's media choices, and how they may contribute to increasing polarization along epistemic lines. At the same time, policy and media literacy interventions should offer more transparency in how algorithmic biases may operate, whilst giving citizens the tools they need to resist such biases, or at least make a well-informed conscious choice on the extent to which their mediated reality is governed by algorithmic biases.

Michael Hameleers

The University of Amsterdam

OTT-Media Selectivity, Algorithmic Personalization, and Audience-User Data: Tailored, Pushed, or Fair?

In the domain of OTT media, we find conflicting viewpoints between doomsayers and enthusiasts of algorithmic personalization. Skeptics quickly predict recurring problems of media business with built-in bias in their recommended contents. But Silicon-Valley enthusiasts easily find reasons to celebrate their algorithms for finding ‘right’ contents that match tastes; thus, sacrificing the privacy of personal data is a justifiable cost of delivering diverse content to unique individuals.

Algorithmic personalization invites debates on more than two sides of pros and cons, however, defying understandings of media audience based on ‘one-to-many’ broadcasting model (Crain, 2018; Guzman & Lewis, 2020). We argue in this forum that taking a side ‘too quickly’ with either skeptics or enthusiasts is a dangerous enterprise that flames moral panic or simply ignores complex challenges posed by OTT platforms. Understanding this complexity, we see obvious missing puzzles in prior debates about practices of OTT-media in their algorithmic content curation.

Missing Puzzle 1: Do Users Follow Preferences?

Here are some facts. According to Netflix’s estimate, an automatic recommendation is responsible for 80% of content selection. On YouTube, as much as 70% of time spent is on watching videos that its algorithm recommends (Cooper, 2021; Thurman et al., 2018). This is startling, as it indicates that the content selection is influenced by the platforms, not by users, telling that preferences are being built algorithmically, rather than users finding the ‘right’ content matching their preferences (cf. information forefending, Kim & Grunig, 2011).

Empirical research on media audiences is not optimistic about human agency. In the earlier broadcasting era, we were inundated by an abundance of agenda-setting research telling us that traditional media audiences are not good at selecting ‘what to think about’ (McCombs, 2005; Proferes & Summers, 2019). The literature on new media audiences quickly piled up to reveal a flimsy side of user agency in regulating their data. Scholars documented: confusion, lack of knowledge, and inadequate data management (Buchi et al., 2017); psychological resignation (Draper & Turow, 2019); technological fatalism (Lutz & Tamò-Larrieux, 2020); and third-person effect (privacy violation only serious to others) (Kim & Hancock, 2016).

The evidence offers us a clue that users’ participation in algorithmic personalization is a product of feeble decisions—namely, a binary choice on whether to opt-in or out of an OTT platform. Put differently, users remain set up algorithmically in a position to be ‘pushed’ for the disclosure of personal data, which is curated to prioritize certain contents (DeVito, 2017; Park, 2021a).

Missing Puzzle 2: Do OTT-Media Platform Follow Preferences?

In a perfectly transparent world, OTT platforms know precisely what user preferences are, thus with their algorithms used to find the best match between content and individuals. But is it the case? This question is seldomly asked concerning the internal validity of algorithmic performance (Sandvig et al., 2016). From the standpoint of managerial perspective, there is no more ethical practice than ‘doing it right’—to deliver the ‘right’ content matching preferences as promised in exchange for personal data. We raise two questions, however.

First, technical infeasibility: Personal data may be the best proxy used to estimate individual preferences. However, whether users will be accurately captured in curated data raises the issue of feasibility, i.e., how feasible it is to recognize and aggregate idiosyncratic tastes into patterns, and rank users according to a programmable schema (Napoli, 2019). Inferences from data introduce bias to certain demographics, and the distance between ‘what is measured of a user’ and ‘what she/he is’ remains technically elusive to overcome (Livingston, 2019; Park, 2021b).

Second, economic infeasibility: Setting aside imperfect algorithmic performance, allocating an infinite number of contents precisely according to diverse tastes in endlessly fragmented markets is not economically viable. Certainly, overselling their capacities to predict and match distinctively unique preferences is a smart business strategy of appealing to advertisers as well as platform subscribers, but a dubious one. Despite a drop in production cost, developing (and maintaining) ‘good’ platform content that achieves economies of scale is an expensive enterprise, involving the risk of being high-sunk with a ‘hit or miss’ (Vonderau, 2015).

Algorithmic Personalization—Tailored, Pushed, or Fair?

The fact that a user is likely to click on algorithmically-assigned entertainment on Netflix or YouTube is not necessarily evidenced that aggregated content is the ‘right’ one. Instead, it may be simply the case of availability bias—users choose what algorithm has made available for immediate choices at moment. What seems to be working in favor of OTT, however, is platform choice, i.e., explicit decisions are to be rendered upon the entry into a platform to access its content. Cognitive psychology literature (Shin et al., 2021; Park, 2022) tells us that when people are presented with limited sets of options, they opt for immediately accessible convenience—preferring instant gratification over the cognitive burden of deliberately sorting out preferences, or even contradicting their expressed interests.

The thesis that we put forth in this forum is that algorithmic personalization is a construct of deliberate choices as OTT platforms set up the conditions with finite choices that proximate user preferences. A profound myth to debunk is that it is possible to detect *what users want* (tastes) and deliver *what they need* (content), as if there were an intrinsic relationship between the accuracy of the personalization and the extent of data collection – the false premise of the validity of algorithmic personalization. We are not denying the growing power of algorithmic personalization. Rather, it is our invitation in this forum to attend to the condition of algorithmic systems in which OTT-media assume unsolicited roles of custodianship by pushing preferences (Gillespie, 2020), amplifying particular voices, contents, and viewpoints that serve platform interests, thus eliminating others. The reason why it is critical to understand this custodianship is that personalization is built upon selectivity, a construct deliberately chosen to link curated data to estimate what users want and need. The specifics of this construction get never negotiated or altered, but being simply assigned to users based on their contribution to the bottom line.

Blind Spots in Algorithmic Fairness Debate

The lesson is that what users *want and need* is algorithmically constructed with multiple options reduced to a handful of offerings in highly selected fashions. We must critically attend to the claim by OTT-media about capacities (of finding and matching preferences), as their algorithmic calculus cannot simply detect and deliver what people want and need, but rather create it. In this regard, the power of the OTT-media platform lies in dictating users about *what to choose from*, not *what to choose, watch, or like* (Couldry & Mejias, 2020; Park, 2021a).

Subsequently, what we posit in this forum is an ethical conundrum, often construed as a rational choice between the need for data surveillance and algorithmic performance, as that choice is built upon the falsified notion that fairness is the function (*F*) of personal data—that is:

- *F* ($\sum$X) = Y, where X = data surveilled and Y = the extent of algorithmic fairness, given greater return of content utility is equal to greater aggregate sum $\sum$ of personal data N (i.e., a cumulation of separately observed incidents, *n*).

Final Notes

One might shrug off to suggest that the current landscape of the digital ecosystem, particularly of OTT-media, is still at its best in the absence of better alternatives. We beg to differ, by offering fruitful lines of considerations for media managers, scholars, and policymakers to ponder about ethical data practice. We propose that the debate should start with the honest acknowledgment that purely mathematical solutions will never reach the optimal point of fairness (Mitchell et al., 2021). We are not expressing this as a vacuous statement, but as an invitation for realistic assessment regarding algorithmic (in)capability in its uninvited custodian roles, which cannot resolve complex ethical objectives by having more data points or different model choices.

Yong Jin Park

Howard University

Jeong Nam Kim

The University of Oklahoma

It's Not (Just) the Algorithm:

Studying the Complex System of Algorithmic Media Curation

Algorithmic media curation reflects a complex socio-technical system that includes human and non-human actors with different motivations, pressures, values, and normative expectations. Any supposed deviation of normative expectations of media distribution, whether that be partisan bias, availability of misinformation, radicalization, and so on, emerges from a complex interplay of actors—algorithms, users, and content creators—and must therefore be evaluated as such. In this essay, we first describe the role of several of the actors in the complex system of algorithmic media. Then, we propose a few research paths to explore the impact of the relationships between actors which help to focus research beyond the analysis of only the algorithms themselves.

Previous work on algorithmic media bias tends to focus on the frequency of the representation of sources or content (Bandy & Diakopoulos, 2021; Muddiman, 2013). Often, that work is specifically about partisan or ideological bias (Diakopoulos et al., 2018; Hu et al., 2018; Kulshrestha et al., 2019; Metaxa et al., 2019; Puschmann, 2018; Robertson et al., 2018), sometimes exploring where bias can come from, including the role of user input (Lurie & Mulligan, 2021; Trielli & Diakopoulos, 2020), and personalization (Bozdag, 2013; Le et al., 2019). We argue for a wider interpretation of curation in algorithmic media, looking not only at the algorithm or the users in isolation, but also at the intersections in curation processes between algorithmic mediators, the actions of the users, and the availability of content provided by publishers or creators.

The actors in algorithmic media

The most prominent actor in algorithmic media is the algorithm itself. Much has been studied about the potentials and harms caused by algorithms in the distribution of media

(Goldman, 2008; Müller et al., 2018). The work that is specific to media has been inspired by wider studies of automated decision-making processes (Diakopoulos, 2015). We know that biases emerge from the algorithms themselves as they make their human-defined editorial choices (Bozdag, 2013; DeVito, 2017). But there are still opportunities to study the degree to which these algorithms are sensitive to and dependent on the input of users and limited by the availability of content.

When it comes to users, research has explored their role in algorithmic bias, using and updating traditional media theories of the 20th century, such as selective exposure theory (Knobloch-Westerwick et al., 2015), and how algorithms reinforce or counteract individual-level bias in media selection (Knobloch-Westerwick et al., 2015; Trielli & Diakopoulos, 2020). Other work has focused on the potential benefits or harms of personalization and customization (Goldman, 2008; Müller et al., 2018) and collaborative filtering, in which media is recommended according to taste predictions (Bozdag, 2013). An under-explored element in algorithmic media is on the supply side: the corpus available for algorithmic curation generated by media creators. As algorithms try to supply users with relevant media, they make that selection on large, but limited media inventories. Some algorithmic media platforms such as streaming services limit inventory based on content licenses. Others, such as search and social media, can restrict what type of media based on curation criteria (e.g. Google reducing visibility of low-quality websites). These criteria can also be conduits for bias, such as if low-quality correlates with political orientation. Content creators, on the other hand, are increasingly aware not only of the needs of the user but the demands of the algorithm as well (Petre, 2021).

As we see, just describing the actors in algorithmic media curation makes it hard to disentangle one from the other: the algorithm is shaped by the values imbued in it by designers, but also by the needs of the user and the opportunities of the source media; the user shapes and responds to the algorithm in search for the available content; the available content is shaped by the needs of the user as refracted through the pressures of the algorithm as understood by the creators. This is the classic description of a complex system (Meadows, 2008), with intricate flows and feedback loops that give rise to the stock of media and information provisioned by the system.

Studying the Complex System

Here, we propose opportunities for applying critical perspectives to the interplay among the actors in these systems, focusing on two main issues: power and values. Power here is each actor's ability to shape the algorithmic curation. This power varies according to the type of platform (i.e. search engines are triggered by the user including keywords; social media platforms by the user choosing to follow other accounts) and by individual platforms (different search engines might have different sensitivities to keywords; different social media platforms might nudge new accounts to follow more often). The power of each actor to shape curatorial bias is ripe for algorithmic accountability investigations (Diakopoulos, 2015) to measure the impact of each actor (Trielli & Diakopoulos, 2020). Such investigations need to be cognizant of system complexity in accounting for responsibilities, including the design orientation of platforms that give or take more power to and from the user. Yet the difficulty of isolating actors (and responsibilities) in the tangled relationships of the complex algorithmic media system underscores the challenge ahead for research.

Second, there is the issue of values: how congruent are the values between the actors in the interaction, and where do they conflict. For example, in platforms that curate news content, there might be a conflict between the journalistic value of the content producer and the value of engagement of the platform. In platforms that are created in the United States and used in other countries, there might be conflicts of local cultures and values in determining important or harmful content. Even commercial values of ad revenue or subscription-oriented content can generate conflict. This type of research on the congruence of values from each actor animates mixed-methods research about different instances of algorithms in media, and the understanding of the motivations for each actor in those instances (e.g., news; entertainment) (DeVito, 2017; Trielli & Diakopoulos, 2019).

From the critical perspectives of power and values, we can explore specific social and technical manifestations of the interplay between actors. One of those manifestations is collusion: what happens if two of the three actors are beholden to the same interests, particularly commercial ones? Another would be colonialism, where global platforms export their values with little room for local dissent.

Complexity as the future of algorithmic media research

Algorithmic media is a complex socio-technical system that includes different actors whose motivations, pressures, and values influence the curation of content. We describe three actors for illustration, but the system is even more complex, including producers such as advertisers or end-users. Future work should delineate more complex relations in this system. Additionally, we describe two lenses to analyze the interplay between actors in algorithmic media systems: power and values. Ultimately, we argue for research incorporating a wider interpretation of curation in algorithmic media, looking at the social and technical intersections of the actors embedded in these complex systems.

Daniel Trielli and Nicholas Diakopoulos,

Northwestern University

Matters of Responsibility for AI in Journalism: Directions for Future Research in the Socio-Technical Study of News

To understand AI, we need to better understand society. In recent academic and policy debates, there has been considerable focus on "responsible" AI—such as building systems that respect professional values and human rights. Similar emphases have been made in conversations about AI in relation to media and journalism (e.g., Broussard et al., 2019). The recent draft of the European Commission for AI Regulation is just one of many examples of initiatives to make AI more human-centric, ethical, and responsible. AI is, in this context, typically understood from a technological venture point as a form of software (see Art. 3(1) of the Draft AI Regulation), automated processing (Art. 22(1), General Data Protection Regulation), or technical system that relies on the analysis of large quantities of data (European Parliament, 2019). Technological systems, however, cannot be more responsible than the users and institutions that adopt them, and the design, operationalization, and functioning of AI-driven applications are influenced by a diverse ecology of agents (Kitchin, 2017; AlgorithmWatch 2020; Diakopoulos 2019) as well as the governance systems that shape their relationships (Van Dijk 2020). This realization contains an uncomfortable truth: It is not enough to make technology more responsible or to design fairer, more diverse, or transparent systems. We need to account for responsibility at the level of society—of the institutions and powers that wield the technology.

Changing the perspective from AI as a technology to AI as a broader socio-technical system opens new avenues for research into responsibility and the way that AI is implicated in the practice and scholarship of journalism. We conceive of responsibility as connected to accountability and the moral and/or legal requirements associated with any agent (e.g., see the "technology paternalism" perspective described in Spiekermann & Pallas, 2006). Building on the

Four A's socio-technical framework put forward by Lewis and Westlund (2015), we approach AI as a socio-technical system that, as a key technological *actant*, intersects with social *actors* that are central to the shaping of media production (i.e., journalists, publishers, regulators, platforms, etc.) as well as diverse *audiences* (i.e., media users, or those on the consumption side of media distribution). These actors, actants, and audiences are interconnected through the *activities* of media work, and it's in the interactions that occur among them—the stuff of which a socio-technical system is made—that we proceed to discuss central developments relating to questions of AI and responsibility, and offer our recommendations for a future research agenda. Given the space limitations of this essay, we focus particularly on a set of key agents—from publishers and platforms, on the one hand, to audiences, regulators, and researchers, on the other—and consider their AI-related roles and responsibilities as well as potential directions for future research.

***Publishers*.** The responsibility for AI among publishers involves the IT/tech department as well as top managers from editorial and business departments. Publishers must focus on systemic solutions taking into account the organizational structures as well as institutional dependencies on third-party AI providers. Publishers use AI for analytics, production of news, personalized news distribution as well as programmatic advertising, and so forth. We call for research focusing on how publishers take organizational responsibility for AI in such activities, involving both proprietary and non-proprietary technology solutions.

***Platforms*.** While there are many platform companies in the world, a small number of these have gained tremendous power on a global scale as key providers and users of AI-driven applications. These players are under growing pressure from governments around the world to make their content moderation and recommender systems more responsible. Doing so cannot be a matter of technological fixes alone. Taking responsibility for their recommendation and content moderation algorithms also requires taking responsibility for those that develop the systems (and ensuring they do so taking into account human rights and public values), those that use the systems (and adopting effective and non-discriminatory ways of enforcing community standards), those that contribute to fixing the systems, such as human content moderators and trusted flaggers (creating fair and humane conditions under which they can do this important work), and those that are affected by their algorithms, including the publishers that depend on them. We call for research that brings together this broader picture of platform responsibility. Additionally, future research would do well to investigate how platforms could employ AI to reduce forms of dark participation (Quandt, 2018), rather than fuel such to drive engagement and advertising revenue.

***Audiences*.** Audiences can be conceived of as recipients of and active participants in news, as well as approached by publishers and advertisers as measurable commodities (Lewis & Westlund, 2015). Additionally, audiences can and should participate in the responsible uptake of AI—for instance, by becoming more familiar with how their engagement with (or non-use of) particular platforms or services, and the tracking and personalization associated with that, contributes to the kind of media experiences they encounter. While it would be unrealistic to expect news consumers to bear the kind of responsibility associated with tech providers or media publishers, there is ample opportunity to enhance ongoing efforts in media literacy to include greater literacy in AI technologies and techniques—thereby giving people more tools in making sense of how information is made and how it moves in the world. We call for research focusing on how audiences' intended as well as unintended media-related activities relate to developing responsible AI in journalism.

Regulators. The current techno-centric focus in policy initiatives such as the Digital Services Act (EU) and related attempts at identifying high-risk AI systems is an important first step in regulation (and in a specific part of the world)—but it only partially achieves the larger goal of fostering public values and responsible technology adoption. Regulating for responsible AI also means regulating for the responsible use, implementation, oversight, control, and contestation of AI systems. We call for research placing AI in context—assessing, for example, how the implementation of AI changes internal distributions of power and responsibility, how structural dependencies to external actors arise, and how governance frameworks are needed to set the rules of the game not only for technology but also for the actors dealing with it.

Researchers. Focusing on AI in its societal context has important implications for research and the training of researchers. An emphasis on AI in its wider societal, institutional, and economic context broadens the research agenda where the Social Sciences/Humanities meets Computer Science (SSH-CS), from the design of fair and responsible algorithms to the design of improved institutions, professional workflows, and social relationships. We call for researchers to study AI in concrete organizational settings for purposes of understanding how AI affects the distribution of responsibilities in the broader network of actors—and, to this end, to consider research collaborations that involve a larger range of disciplines and stakeholders involved.

Natali Helberger

The University of Amsterdam

Seth C. Lewis

The University of Oregon

Oscar Westlund

Oslo Metropolitan University

The Bias in Media regarding Algorithmic Bias of AI

Applications of artificially intelligent (AI) algorithms are becoming increasingly common in many realms of life to assist with making or even fully automating decisions. AI technologies currently suffer from the hype surrounding them, according to which their potential borders on utopian, while at the same time they are being fiercely criticized for being culturally biased, unfair, unethical, or downright evil. While media have become deeply engaged in the debate, especially supporting critical positions (Kaplan & Haenlein, 2020; Müller et al., 2018), there are still few academic studies of the phenomenon in media and journalism research (Baumann, 2020). Why are we seeing such a sharp split in the arguments around AI in the public discussion but, simultaneously, a reluctance to engage in the debate in the scholarly sphere? This contribution aims at shedding light on AI by clarifying what it actually is and hence doing away with some of the misconceptions underlying the discussion. It will also outline ideas about the beneficial use of AI technologies as presented in media and journalism, as well as describe some limitations and challenges.

Let's start with some misconceptions. A common misunderstanding is that AI is a single technology, when in fact AI refers to a broad array of technologies; Machine Learning is currently the most prominently applied type. AI technologies have in common that they mimic human intelligence as they interpret external data and use what they learn to flexibly adapt their behavior as they try to achieve specific goals and tasks (Guzman & Lewis, 2020; Kaplan & Haenlein, 2020; Meske et al., 2022). AI technologies are categorized into evolutionary stages—narrow, general and super-intelligent—and are often classified as analytical, human-inspired, or humanized depending on their cognitive, emotional, and social competencies. The discussion of

AI often mixes up these different stages and types, which creates confusion about the term “AI” itself (Chan-Olmsted, 2019).

The uses of AI technologies are manifold, touching on all realms of life. Prominent examples include their use or potential use in self-driving cars, health monitoring systems, recruitment tools, and of course military equipment. Potentially adverse applications and undesirable results of AI systems, such as accidents, and automated recommendations that are clearly biased typically feature in the media. Negative examples attract a larger audience. Positive examples of AI applications are underrepresented in current reporting on AI technologies, and so it’s fair to state that the media carry a negative bias against AI.

Let’s investigate bias against AI in more detail. AI technologies are being criticized for having an algorithmic bias, which means that supposedly malicious algorithms are at the heart of the problems with AI. In media discourse, the term “algorithm” has almost become synonymous with something we cannot control, almost as if algorithms have lives of their own. But what exactly is an algorithm? It is nothing more than a unique set of instructions for solving a problem (or a class of problems) or for achieving a goal. A cooking recipe is an algorithm. It tells us which ingredients we need in what quantities and describes the steps to follow to create a final dish. Thus, algorithms consist of a finite number of well-defined individual steps. These steps can be formulated in human languages, as in a recipe, or implemented in an executable computer program. While people can commonly read a recipe and follow the instructions, programming languages are typically unfamiliar. However, they are simply a way of formulating instructions, just like sheet music captures the instructions for playing a certain musical piece. Instructions can carry bias if the human who formulates them intentionally or unintentionally ingrains bias into the recipe or computer program. Thus, it is not the algorithm that is biased, but the human creator.

AI technologies have also been criticized for producing biased results and recommendations (e.g. Ananny & Crawford, 2018; Hoffman, 2019). An overly cited example is that of AI systems engaged in job selection that favor certain genders or ethnicities. The alleged conclusion is that AI technologies invariably produce biased outputs. To explain what happened in these cases, two important variables need to be considered: the input data and how the training of a machine learning type AI system works. The main difference between AI and traditional solution-seeking methods is that the latter uses a given method to transform input data into (unknown) output data, while in AI existing data is used to train the system, which then finds the "method" to produce expected outputs (Chivers, 2021; Moore et al., 2021). The outputs are assessed by humans and their feedback helps the AI system further improve the "method." Bias can occur if the training data contains an inherent bias, e.g., that in the past recruiters favored particular age groups or genders. Likewise, bias can occur during the output assessment if the humans involved (unconsciously or consciously) transfer their own bias to the system. In both cases, the bias is human-made and should not be blamed on the AI. In fact, the AI system merely holds up a mirror that reflects the bias already inherent in the setting. And if this leads to making implicit bias explicit, the ensuing debate can help to pave the way to reducing or even eliminating such bias.

Another issue that is commonly raised against AI is that it is unknown how an AI system operates internally, especially given that the learning process of such a system creates non-transparent dependencies between input data and outputs (Meske et al., 2022). This Black Box phenomenon has raised calls for "explainable AI," explainable in the sense that it enables human users to appropriately understand and effectively manage the systems as a prerequisite of a trustworthy and responsible AI (Adadi & Berrada, 2018; European Commission & Directorate-General for Communications Networks, Content and Technology, 2019; Thiebes et al., 2021).

Applications of AI systems in media and journalism can have many benefits. AI systems can support humans in cumbersome tasks such as analyzing and organizing large data sets (Hartmann et al., 2019) and in making good decisions. For specialized cases such as, for example, creating metadata, they already reach or even surpass the task performance of humans. In customer interaction, they play a major role by recommending engaging content and allowing for targeted information distribution (Baumann, 2021; Chen et al., 2019). Nevertheless, super-intelligent general AI systems are still far in the future and the purported danger that AI systems will eventually become fully human-like will not happen any time soon. Judgment used by AI systems relies on trained reasoning, but humans can also use imagination, reflection, valuation, and empathy. We should not forget our moral responsibilities—and that includes avoiding bias in reporting on AI technologies. Otherwise we may miss out on harnessing AI technologies for the benefits they can bring.

Sabine Baumann

Jade University of Applied Sciences

References

Adadi, A., & Berrada, M. (2018). Peeking Inside the Black-Box. *IEEE Access, 6,* 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052

AlgorithmWatch and Bertelsmann Stiftung. 2020. “Automating Society Report 2020.” Automating Society Report 2020. 2020. automatingsociety.algorithmwatch.org

Ananny, M., & Crawford, K. (2018). Seeing without knowing. *New Media & Society, 20*(3), 973-989. https://doi:10.1177/1461444816676645

Arendt, H. (1967, February 25). *Truth and Politics*. The New Yorker

Bennett, L. W., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication, 33*(2), 122-139. https://doi.org/10.1177/0267323118760317

Borgesius, F., Trilling, D., Möller, J., Bodó, B., Vreese, C., & Helberger, N. (2016). Should we worry about filter bubbles?. *Internet Policy Review*. Retrieved from https://policyreview.info/articles/analysis/should-we-worry-about-filter-bubbles

Baek, Y. M., Kim, E. M., & Bae, Y. (2014). My privacy is okay, but theirs is endangered. *Computers in Human Behavior, 31*, 48-56.

Bandy, J., & Diakopoulos, N. (2021). Curating Quality? How Twitter’s Timeline Algorithm Treats Different Types of News. *Social Media + Society, 7*(3), 20563051211041648.

Baumann, S. (2021). Media Management and Business: Old vs. New. In Mahoney, M. & Tang, T. (Eds) (2021) *The Handbook of Media Management and Business* (pp. 407-425). Lanham: Rowman & Littlefield.

Baumann, S. (2020). Guest editor’s introduction: strategic media management at a junction, *Journal of Media Business Studies, 17(1),* 1-12. DOI: 10.1080/16522354.2020.1740564

Büchi, M., Just, N., & Latzer, M. (2017). Caring is not enough: the importance of Internet skills for online privacy protection. *Information, Communication & Society, 20*(8), 1261-1278.

Bozdag, E. (2013). Bias in algorithmic filtering and personalization. *Ethics and information technology,* 15(3), 209-227.

Broussard, M., Diakopoulos, N., Guzman, A. L., Abebe, R., Dupagne, M., & Chuan, C.-H. (2019). Artificial Intelligence and Journalism. *Journalism & Mass Communication Quarterly*, *96*(3), 673–695.

Chan-Olmsted, S. M. (2019). A review of artificial intelligence adoptions in the media industry. *International Journal on Media Management, 21*(3-4), 193–215. https://doi.org/10.1080/14241277.2019.1695619

Chen, G., Xie, P., Dong, J., & Wang, T. (2019). Understanding Programmatic Creative. *Journal of Advertising, 48(*4), 347–355. https://doi.org/10.1080/00913367.2019.1654421

Chivers, T. (2021). How to train an all-purpose robot: DeepMind is tackling one of the hardest problems for AI. *IEEE Spectrum*, 35–41. https://spectrum.ieee.org/how-deepmind-is-reinventing-the-robot

Crain, M. (2018). The limits of transparency. *New Media & Society, 20*(1), 88-104. https://doi:10.1177/1461444816657096

Cooper, P. (2021). How Does the YouTube Algorithm Work in 2021? The Complete Guide. Available at: https://blog.hootsuite.com/how-the-youtube-algorithm-works/

Couldry, N., & Mejias, U. A. (2020). The costs of connection: how data are colonizing human life and appropriating it for capitalism.

DeVito, M. A. (2017). From editors to algorithms. *Digital Journalism*, *5*(6), 753-773.

Diakopoulos, N. (2019). *Automating the News: How Algorithms Are Rewriting the Media*. Harvard University Press.

Dijck, J. van. (2020). Governing digital societies: private platforms, public values. *Computer Law & Security Review* 36 (April).

Diakopoulos, N. (2015). Algorithmic accountability. *Digital journalism, 3*(3), 398-415.

Diakopoulos, N., Trielli, D., Stark, J., & Mussenden, S. (2018). I vote for—how search informs our choice of candidate. Digital Dominance: The Power of Google, Amazon, Facebook, and Apple, M. Moore and D. Tambini (Eds.), 22.

Dörr, K., & Hollnbuchner, K. (2017). Ethical challenges of algorithmic journalism. *Digital Journalism, 5*(4), 404-419. doi:10.1080/21670811.2016.1167612

Draper, N. A., & Turow, J. (2019). The corporate cultivation of digital resignation. *New Media & Society, 21*(8), 1824-1839.

European Commission, & Directorate-General for Communications Networks, Content and Technology. (2019). Ethics guidelines for trustworthy AI. Publications Office. https://doi.org/10.2759/177365

European Parliament, Directorate-General for Parliamentary Research Services. (2019). Understanding Algorithmic Decision-Making: Opportunities and Challenges. LU: Publications Office. https://www.europarl.europa.eu/stoa/en/document/EPRS_STU(2019)624261

Gillespie, T. (2020). Content moderation, AI, and the question of scale. *Big Data & Society, 7*(2), 2053951720943234.

Guzman, A. L., & Lewis, S. C. (2020). Artificial intelligence and communication. *New Media & Society, 22*(1), 70-86.

Hartmann, J., Huppertz, J., Schamp, C., & Heitmann, M. (2019). Comparing automated text classification methods. *International Journal of Research in Marketing*, *36*(1), 20–38. https://doi.org/10.1016/j.ijresmar.2018.09.009

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

Hoffmann, A. (2019). Where fairness fails. *Information, Communication & Society, 22*(7), 900-915. https://doi:10.1080/1369118X.2019.1573912

Goldman, E. (2008). *Search engine bias and the demise of search engine Utopianism web search* (pp. 121–133). Berlin Heidelberg: Springer.

Hu, D. S., Jiang, S., Robertson, R. E., & Wilson, C. (2019). Auditing the partisanship of Google search snippets. In Web conference 2019: *Proceedings of the world wide web conference* (pp. 693–704). https://doi-org.turing.library.northwestern.edu/10.1145/3308558.3313654

Kaplan, A., & Haenlein, M. (2020). Rulers of the world, unite! *Business Horizons, 63*(1), 37–50. https://doi.org/10.1016/j.bushor.2019.09.003

Kitchin, R. (2017). Thinking critically about and researching algorithms. *Information, Communication & Society* 20(1): 14–29.

Kim, S. J., & Hancock, J. T. (2015). Optimistic bias and Facebook use. *Cyberpsychology, Behavior, and Social Networking, 18*(4), 214-220.

Kim, J.-N., & Grunig, J. E. (2011). Problem solving and communicative action. *Journal of Communication, 61*, 120-149.

Knobloch-Westerwick, S., Johnson, B. K., & Westerwick, A. (2015). Confirmation bias in online searches. *Journal of Computer-Mediated Communication, 20*(2), 171–187.

Kulshrestha, J., Eslami, M., Messias, J., Zafar, M. B., Ghosh, S., Gummadi, K. P., & Karahalios, K. (2019). Search bias quantification: Investigating political bias in social media and web search. *Information Retrieval Journal, 22*(1-2), 188–227.

Le, H., Maragh, R., Ekdale, B., High, A., Havens, T., & Shafiq, Z. (2019, May). Measuring political personalization of Google news search. In The World Wide Web Conference (pp. 2957-2963). http://dx.doi.org/10.1145/3308558.3313682

Lewis, S. C., & Westlund, O. (2015). Actors, actants, audiences, and activities in cross-media news work. *Digital Journalism*, *3*(1), 19–37.

Lurie, E., & Mulligan, D. K. (2021). Searching for Representation: A sociotechnical audit of googling for members of US Congress. arXiv preprint arXiv:2109.07012.

Livingstone, S. (2019). Audiences in an age of datafication. *Television & New Media*, *20*(2), 170-183.

Lutz, C., & Tamò-Larrieux, A. (2021). Do privacy concerns about social robots affect use intentions? *Frontiers in Robotics and AI, 8,* 63.

Marwick, A., & Lewis, R. (2017). Media manipulation and disinformation online (pp. 1-104). *Data and Society Research Institute*. Retrieved September 18, 2018, from https://datasociety.net/output/media-manipulation-and-disinfo-online/

Meske, C., Bunde, E., Schneider, J., & Gersch, M. (2022). Explainable artificial intelligence. *Information Systems Management, 39*(1), 53-63. https://doi.org/10.1080/10580530.2020.1849465

Möller, J., Trilling, D., Helberger, N., & van Es, B. (2018). Do not blame it on the algorithm. *Information, Communication & Society, 21*(7), 959-977. doi:10.1080/1369118X.2018.1444076

McCombs, M. (2005). A look at agenda-setting. *Journalism studies, 6*(4), 543-557.

Meadows, D. H. (2008). *Thinking in systems: A primer.* Chelsea Green Publishing.

Metaxa, D., Park, J. S., Landay, J. A., & Hancock, J. (2019). Search media and elections. *Proceedings of the ACM on Human-Computer Interaction,* 3(CSCW), 1–17. https://doi.org/10.1145/3359231

Mitchell, S., Potash, E., Barocas, S., D'Amour, A., & Lum, K. (2018). Prediction-based decisions and fairness. arXiv preprint arXiv:1811.07867.

Muddiman, A. (2013). Searching for the next U.S. president. *Journal of Information Technology & Politics, 10*(2), 138–157. https://doi.org/10.1080/19331681.2012.707440

Müller, J., Trilling, D., Helberger, N., & van Es, B. (2018). Do not blame it on the algorithm. *Information, Communication & Society, 21*(7), 959–977. https://doi.org/10.1080/1369118X.2018.1444076

Moore, S. K., Schneider, D., & Strickland, E. (2021). How Deep Learning Works. *IEEE Spectrum* (October), 32–33. https://spectrum.ieee.org/what-is-deep-learning

Napoli, P. M. (2019). *Social media and the public interest.* Columbia University Press.

Park, Y. (2019). Role of fairness, accountability, and transparency in algorithmic affordance. *Computers in Human Behavior, 98*, 277-284. http://doi.10.1016/j.chb.2019.04.019

Park, Y. J. (2021a). *The future of digital surveillance*. University of Michigan Press.

Park, Y. J. (2021b). Structural logic of ai surveillance and its normalisation in the public sphere. *Javnost-The Public,* 1-17. https://doi.org/10.1080/13183222.2021.1955323

Park, Y. J. (2022). Personal data concern, behavioral puzzle and uncertainty in the age of digital surveillance. *Telematics and Informatics, 66,* 101748.

https://doi.org/10.1016/j.tele.2021.101748

Petre, C. (2021). *All the news that's fit to click: How metrics are transforming the work of journalists*. Princeton University Press.

Proferes, N., & Summers, E. (2019). Algorithms and agenda-setting in Wikileaks'# *Information, Communication & Society, 22*(11), 1630-1645.

Puschmann, C. (2018). Beyond the bubble: Assessing the diversity of political search results. *Digital Journalism, 7*(6), 824–843. https://doi.org/10.1080/21670811.2018.1539626

Quandt, T. (2018). Dark participation. *Media and communication*, *6*(4), 36-48.

Poon, M. (2016). Corporate capitalism and the growing power of big data. *Science, Technology, & Human Values, 41*(6), 1088-1108.

Robertson, R. E., Jiang, S., Joseph, K., Friedland, L., Lazer, D., & Wilson, C. (2018). Auditing partisan audience bias within google search. *Proceedings of the ACM on Human-Computer Interaction,* 2(CSCW), 1-22.

Sandvig, C., Hamilton, K., Karahalios, K., & Langbort, C. (2016). Automation, algorithms, and politics| when the algorithm itself is a racist. *International Journal of Communication, 10*, 19.

Shin, D., Rasul, A., & Fotiadis, A. (2021). Why am I seeing this? *Internet Research.* https://doi.org/10.1108/INTR-02-2021-0087

Shin, D. (2021). The perception of humanness in conversational journalism. *New Media and Society.* http://doi:10.1177/1461444821993801

Starbird, K. (2019). Disinformation's spread. *Nature*, *571*(7766), 449.

Spiekermann, S., & Pallas, F. (2006). Technology paternalism–wider implications of ubiquitous computing. *Poiesis & praxis*, *4*(1), 6-18.

Thiebes, S., Lins, S., & Sunyaev, A. (2021). Trustworthy artificial intelligence. *Electronic Markets, 31*(2), 447–464. https://doi.org/10.1007/s12525-020-00441-4

Thurman, N., & Fletcher, R. (2018). Are newspapers heading toward post-print obscurity? *Digital journalism*, *6*(8), 1003-1017.

Thurman, N., Moeller, J., Helberger, N., & Trilling, D. (2019). My friends, editors, algorithms, and I. *Digital Journalism*, *7*(4), 447-469. https://doi.org/10.1080/21670811.2018.1493936

Trielli, D., & Diakopoulos, N. (2019, May). Search as news curator: The role of Google in shaping attention to news information. In *Proceedings of the 2019 CHI Conference on human factors in computing systems* (pp. 1-15).

Trielli, D., & Diakopoulos, N. (2020). Partisan search behavior and Google results in the 2018 US midterm elections. *Information, Communication & Society*, 1-17.

Vonderau, P. (2015). The politics of content aggregation. *Television & New Media, 16*(8), 717-733.

Waisbord, S. (2018). Truth is what happens to news. *Journalism Studies, 19*(13), 1866-1878.

Wölker, A., & Powell, T. (2021). Algorithms in the newsroom? *Journalism*, *22*(1), 86-103. https://doi.org/10.1177/1464884918757072

Zhang, Y., Lukito, J., Su, M-H., Suk, J., Xia, Y., Kim, S. J.,, Doroshenko, L., & Wells, C. (2021). Assembling the networks and audiences of disinformation. *Journal of Communication*. https://doi.org/10.1093/joc/jqaa042

Zimmermann, F., & Kohring, M. (2020). Mistrust, disinforming news, and vote choice. *Political Communication, 37*(2), 215-237. https://doi.org/10.1080/10584609.2019.1686095