

Article

## Post-Publication Gatekeeping Factors and Practices: Data, Platforms, and Regulations in News Work

Margareta Salonen \*, Veera Ehrlén, Minna Koivula, and Karoliina Talvitie-Lamberg

Department of Language and Communication Studies, University of Jyväskylä, Finland

\* Corresponding author ([margareta.salonen@jyu.fi](mailto:margareta.salonen@jyu.fi))

Submitted: 14 November 2022 | Accepted: 6 March 2023 | Published: 28 June 2023

### Abstract

The gatekeeping literature has turned to look at the factors and practices that shape gatekeeping in the post-publication environment, i.e., after news has entered circulation. This article adds to the discussion and argues that news workers share gatekeeping power in the post-publication environment with audiences, platforms, and regulations. Further, this study extends the post-publication gatekeeping framework and considers it in the context of datafication. The article aims to broadly understand how (audience) data is part of editorial decision-making in news media from news workers' perceptions. The current study was conducted by interviewing news workers from three Finnish news organisations. The interview data was analysed utilising qualitative iterative content analysis. Our analysis revealed that the use of (audience) data in news organisations increasingly shapes news workers' journalistic decision-making processes. We found that news workers were ambivalent toward data (use) and that their reliance on platform data depended on the particular platform. Furthermore, when interviewed about journalism ethics, news workers only connected it with legislative issues, such as General Data Protection Regulation. Lastly, we could see that regulatory factors of data, i.e., legislation and media self-regulation, have power over news production and distribution. This study reflects how journalism (research) is shifting from an audience-centric view to a data-driven one, i.e., it is experiencing a *data turn*.

### Keywords

audiences; datafication; gatekeeping; news workers; platforms; regulations

### Issue

This article is part of the issue “A Datafied Society: Data Power, Infrastructures, and Regulations” edited by Raul Ferrer-Conill (University of Stavanger / Karlstad University), Helle Sjøvaag (University of Stavanger), and Ragnhild Kr. Olsen (Oslo Metropolitan University).

© 2023 by the author(s); licensee Cogitatio Press (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

### 1. Introduction

The growing use of audience data in newsrooms has impacted the power dynamics between news organisations, platforms, and audiences. The “digital trinity” of datafication, algorithmisation, and platformisation (Latzer, 2021) has led to a situation where journalists hold less gatekeeping power, i.e., a growing number of players have gained influence over news distribution, circulation, and the business itself (Salonen et al., 2022; Seuri & Ikäheimo, 2022). Through their digital footprints, such as interacting with news on social media, audiences are influencing editorial decision-making (Tandoc

& Vos, 2016), and, in turn, social media platforms' algorithms tangle with news distribution and visibility, affecting the way news is shown to audiences (van Dijck et al., 2018). Algorithmic platforms and their users have entered the news ecosystem and currently intertwine with news processes, especially in the post-publication context, i.e., after news has entered circulation (Hermida, 2020; Salonen et al., 2022).

In this study, we examine datafied news work from the perspective of post-publication gatekeeping. Our aim is to broadly understand how (audience) data is part of editorial decision-making in news media from news workers' perceptions. More specifically, we contribute

to the literature by discussing what role regulations play in the datafied news environment. From the viewpoint of media accountability, news media consider and reflect the external control (such as laws of the country/region), the internal control (such as journalistic values, ethics, and press councils) as well as audiences and other members of civil society (Eberwein et al., 2019) when deciding what is newsworthy to publish. In the current study, we focus on media regulation from a twofold perspective: Firstly, legislation in the European Union, for example, the General Data Protection Regulation (GDPR), ensures and controls users, media organisations, and platform companies' rights over data (Meier & Trappel, 2022); and secondly, news organisations consider media self-regulation in editorial decision-making processes. This study addresses the suggestion of Porlezza and Eberwein (2022), Seuri and Ikäheimo (2022), and Seuri et al. (2022) that media regulation (the external control) and self-regulation (the internal control) should be investigated in the era of datafication.

Datafication is, in this study, understood as digital media's capacity to turn all human action and interaction into measurable digital traces (Breiter & Hepp, 2018). Essentially, knowledge of user characteristics and behaviours has become the main currency in the current media ecosystem (Ohlsson & Facht, 2017): In the process of datafication, (individual) data is monetised for its business potential (Coudry & Yu, 2018). Power over audience data is exercised by those with access to databases and expertise in processing and data mining (Andrejevic, 2014), such as news media and platform companies. Through digital profiling, platform companies can define which content users are exposed to and predict their (future) behaviour. News media depend on these infrastructural services as platform companies have developed considerable market and gatekeeping power (Meier & Trappel, 2022).

Ownership and control over news-related audience data have become complex, and therefore, it is important to understand how news media utilise and make decisions concerning data. In this research, we highlight the (potential) role of audience data from self and/or third-party-governed platforms in news workers' editorial decision-making. We do this by reflecting news workers' perceptions of their datafied working practices through the lens of post-publication gatekeeping theory, which looks at the factors that shape news after its publication. The current study adds to previous research on post-publication factors, practices, and the environment itself (Hermida, 2020; Salonen et al., 2022). Further, this study connects regulatory factors (legislation and media self-regulation) theoretically to the post-publication gatekeeping framework (Hermida, 2020) and provides evidence by applying the suggested extended framework empirically. Finally, this study takes a stance on the well-needed discussion of regulatory factors in the era of datafication.

## 2. Theoretical Framework of (Post-Publication) Gatekeeping in the Context of Datafied News Environment

In the 2020s, the context where news circulates is increasingly digital and datafied. Datafication of news has meant, for example, that newsrooms' editorial decisions are increasingly data-oriented (Vu, 2014). Datafication has also impacted economic models in journalism and challenged journalistic autonomy (Hanusch, 2017). This has led to the datafication of the digital news environment, and hence, the gatekeeping processes have also changed. Gatekeeping has been defined as "the process of culling and crafting countless bits of information into the limited number of messages that reach people each day" (Shoemaker & Vos, 2009, p. 1). The definition still holds up even though the theory has transformed considerably since its birth in the 1940s (Vos, 2019). This section first introduces gatekeeping theory in the digital age, then reviews studies focused on post-publication gatekeeping, and finally, discusses the role of regulations as a post-publication gatekeeping factor in datafied news work.

The digital news environment has brought changes to the ways gatekeeping theory has been utilised and developed for research in the past. In 1989, Abbott and Brassfield (1989) compared print and electronic media, and in 1990, Berkowitz (1990) studied the gatekeeping processes of local television news. More recently, the theory has been applied to online contexts. These studies have looked at the rise of news events on social media (Meraz & Papacharissi, 2013), user-generated visibility on media websites (Singer, 2014), visual gatekeeping practices (Pantti, 2015), and social media editors' impact on news diffusion (Welbers & Opgenhaffen, 2018), for example. Gatekeeping has also been researched from the viewpoint of data and metrics. Tandoc (2014) pointed out how web analytics is changing the gatekeeping process after decades of journalists paying little attention to audience opinion in their decision-making. Consequently, information gained from (audience) data is part of the journalistic gatekeeping process. Nowadays, the question is not about whether data affects newsroom decision-making but, rather, how much. Further, the theory field has started to turn towards the post-publication viewpoint, i.e., the context where news is constantly circulated in the datafied digital news environment.

Some previous studies have looked at the post-publication side of gatekeeping, even though they do not specifically talk about post-publication gatekeeping. Singer (2014) introduced the term secondary gatekeeping, which refers to how users can up or downgrade the visibility of an online news item, while Bruns (2018) talks about how users and journalists can act as gatewatchers to/of information that is relevant to be distributed further online. Wallace (2018), in turn, points out different types of digital gatekeepers in contemporary society: journalists, individual amateurs, strategic professionals,

and algorithms. These studies can be categorised as post-publication gatekeeping studies because they look at the gatekeeping factors (e.g., audiences/users and platforms) and practices (e.g., users' ability to up or downgrade or gatewatch) in the environment after news has been published.

Post-publication gatekeeping specifically has been previously researched by Hermida (2020) and Salonen et al. (2022). Salonen et al. (2022) introduced the concept of conversational gatekeeping, highlighting the social interactional nature of gatekeeping in the post-publication context. Further, they see gatekeeping materialising as and in social interaction: Journalists and audiences negotiate and create the conversational norms together and decide on the accepted content for the particular online platform. Further, Hermida (2020) introduced the framework of four factors (4Ps)—publics, platforms, paraphernalia, and practices—through which post-publication gatekeeping can be viewed and which shape the processes of gatekeeping. Publics refers to the news audiences, reaching from the members of the public to politicians, businesspeople, and journalists themselves. Platforms concern the platform giants such as Google, Meta, and Twitter. Paraphernalia refers to the materiality of gatekeeping, such as mobile devices, smart speakers, and software. Practices refers to social (spatial and temporal) practices around how users engage with the news; for example, whether the news is consumed on a bus or while lying in bed.

In their study, Hermida (2020, p. 16) calls for studies that consider digital metrics' influence on editorial decision-making by considering "how all or some of the four Ps impact flows of news and information post-publication." Specifically, regulation issues related to gatekeeping are under-researched even though regulation increasingly shapes the formation of news. With the term regulations, we refer to the legislative factors, i.e., laws of the country and/or region, and to the self-regulatory factors, namely, journalistic values and ethics. Academics have called for collective regulation of data-driven systems (Steedman et al., 2020). In the EU, legislation (the GDPR and several digital acts, for example) ensures and controls users, media companies, and platform companies' rights over data (Meier & Trappel, 2022). For example, some European data protection authorities have found the use of US-based Google Analytics unlawful due to the increased risk of being in breach of the GDPR that regulates the use of personal data (Roosa et al., 2022). These new digital laws will challenge platform companies' data monopolies, and as a result, companies such as Meta have threatened to withdraw from the EU market (Burgess, 2022). The new regulations will also affect data collected by third parties and used by the media. As a result, audiences may gain more control over their data, and the monetisation model of platforms and media companies could be undermined.

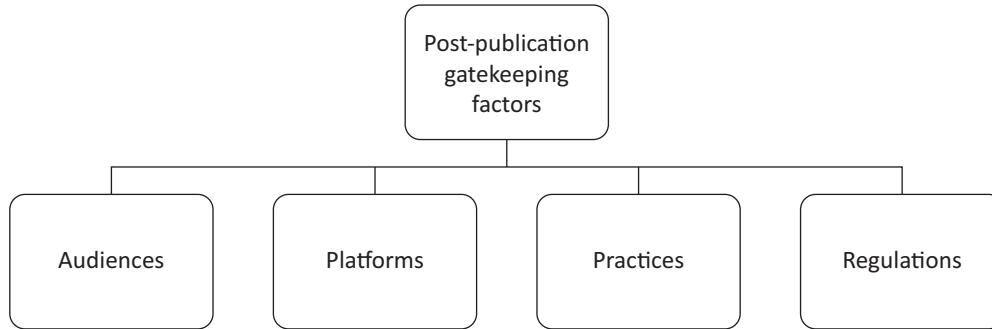
Further, Seuri et al. (2022) discuss the new gatekeeping regime and highlight the role and need to regulate

platform giants and big media companies. In their future scenarios for the platform society, Seuri et al. (2022) see that regulation is needed to gain positive outcomes for the information environment and to counterbalance the network effects that create platform monopolies. Furthermore, Aral (2020) has called for structural reform: Data flows between platforms should be enforced to ensure sustainable interoperability between platforms. This means that user data and actions should no longer be monopolised by platform companies. The Digital Services Act is designed to combat this in the EU, which could mean that in the future, news companies will have better access to databases collected by platform companies.

From the viewpoint of media accountability and journalism ethics, news media are currently trying to balance the journalistic ethos of reporting what the audience needs to know and the data-driven view of what the audience wants to know (Hanusch, 2017). This notion raises ethical questions as the journalistic field also has its norms and ethics on which editorial decisions are based. In Finland, the context of our empirical study, the Finnish press council governs journalistic (self-regulation) guidelines and processes complaints related to them. In relation to media accountability, the press council is an important part of the journalistic institution and has been designed to oversee the media's responsibility to society, citizens, and the journalistic institution itself (Eberwein et al., 2019). However, in digital journalism driven by datafication and algorithms, there is still a lack of normative standards and regulations across Europe (Porlezza & Eberwein, 2022). Journalism scholars (e.g., Porlezza & Eberwein, 2022; Rydenfelt et al., 2022) have argued that journalistic self-regulation needs to be adapted to the era of datafication, news automation, and personalisation.

From these theoretical premises, we answer Hermida's call to apply the framework of the 4Ps (publics, platforms, paraphernalia, and practices) and tackle the question of digital metrics' connectedness to editorial decision-making. Thus, we propose the framework of "Post-publication gatekeeping factors in datafied news work" (Figure 1) that has been built on the premises of the 4Ps and the previous theoretical discussion of regulations and datafication of journalism. In the extended framework, we add the factor of regulations, i.e., legislation and self-regulation, to extend the 4Ps framework and see that all this is taking place in the context of a datafied news environment.

In this study, we use the term audiences instead of publics as it is more descriptive in depicting the people for whom the news is created. Further, we see that the factor of platforms includes all kinds of platforms used by the news media, self-governed and third-party governed. Furthermore, we see that the factor of paraphernalia—the materiality of the object—is omnipresent in the factor of platforms, as platforms are technological constellations. Therefore, this study considers paraphernalia (e.g., software) as embedded in the platform factor.



**Figure 1.** Post-publication gatekeeping factors in datafied news work.

By practices, we refer to the concept beyond audiences’ social practices and include all kinds of journalistic practices that news workers can perform during their daily working routines. Thus, in this study, we have adapted from Hermida’s (2020) study, the factors of audiences, platforms and practices, and added the factor of regulations.

Against this background, this study asks:

RQ: How do news workers perceive the datafied factors of audiences, platforms, and regulations, and how do these factors shape their working practices from the viewpoint of post-publication gatekeeping?

**3. Methods and Materials**

The material for the study was gathered by interviewing Finnish news workers (N = 9). Finland makes for a particularly interesting research context concerning audience data use because news reach and trust in news media are generally high (Newman et al., 2022). In global comparison, Finns trust news sites to use their personal data fairly more than people in other Western countries (Newman et al., 2022). This context may also be reflected in news workers’ attitudes toward data use. We adopted a qualitative approach to comprehend news workers’ perceptions and understandings of data-related issues and practices and interviewed nine news workers from three Finnish newspaper organisations. The first organi-

sation is a large national daily newspaper, and the other two are sizable regional newspapers, both publishing the largest newspaper in their geographical area in terms of circulation.

Our study participants are personnel who deal with data-related questions in their everyday work in news organisations. Their views on data use in news organisations vary somewhat based on their titles, as some worked in content production roles (e.g., social media manager) while others had more managerial tasks (e.g., head of technological development). Importantly, however, all were experts in how their organisation utilises audience data. Hence, we refer to them collectively as news workers. Seven participants identified as male, one as female, and one as non-binary. Participants’ job titles and years of experience can be found in Table 1.

The interviewees were selected for the study by peer recommendations within the organisations. Prior to the interviews, the participants were informed about the research project and the interview themes. They were also asked to review and approve a consent form guaranteeing the voluntary and anonymous nature of participation and the confidentiality of the interview material. Semi-structured interviews were conducted by the first author via Zoom and face-to-face between May and August of 2022. On average, interviews lasted 96 minutes (a total of 866 minutes). The interview themes were: (a) Collection and use of data from news media sites, (b) collection and use of data from social media

**Table 1.** Interviewees’ job titles and work experience in years.

Interviewee	Work title	Years of experience
NW1	Producer	15
NW2	Manager	20
NW3	Lead Developer	2,5
NW4	Social Media Producer	6
NW5	Head of Business Development	8
NW6	Web Manager	12
NW7	Web Analyst	20
NW8	Head of News	13
NW9	Head of Technological Development	10

platforms, (c) ethics and responsibility in data collection and use, (d) news media's approach to social media platforms, (e) news media's approach to legislation dealing with news content and data practices, and (f) moderation of news content and discussions.

The interview material was analysed using qualitative iterative content analysis (Tracy, 2018). This method combines and alternates between deductive and inductive research approaches, i.e., it iteratively combines data and previous theory. This method was chosen because it "focuses on more narrow aspects of the data that have potential to extend specific theories or address practical problems" (Tracy, 2018, p. 63). In this study, the post-publication gatekeeping theory was used to gain an understanding of the current data-related practices and perceptions of news workers. At the same time, the qualitative data provided empirical evidence to support the proposed extension of Hermida's (2020) theory. Further, the iterative approach brought forward news workers' practices related to data use.

The interview material was analysed and coded by the first author. In the first round of analysis, the material was read and viewed in light of *what* was present. This reading highlighted the salience of the post-publication gatekeeping environment. Thus, throughout the following rounds of coding, the interviews were reflected in the light of the post-publication gatekeeping theory to ensure that all the factors based on previous literature (audiences, platforms, and regulations) were taken into account in the analysis. In addition, to ensure that all material discussing data-related practices was thoroughly reviewed, the parts mentioning data generally were coded as a fifth category. By doing so, the authors gained a comprehensive understanding of the interview material and advanced their conception of how post-publication gatekeeping factors are connected to data use in news media. In the final phase of the analysis, the authors jointly evaluated the fit between the formulated categories and coded content through a discussion on data excerpts relating to post-publication gatekeeping factors and different kinds of data-related practices evident in the interview data.

#### 4. Findings: Post-Publication Gatekeeping Factors and Practices

In this section, we examine our data through the lenses of post-publication gatekeeping factors and practices in datafied news work, as Hermida (2020) suggested. We apply the extended framework previously suggested in the theory section (see Figure 1) to our empirical data. In the following, we analyse and provide empirical evidence of how news workers perceive the datafied factors of audiences, platforms, and regulations and how these factors may shape their working practices. First, we explain the significant and ambivalent role of audience data in journalistic decision-making processes; second, we open up the news workers' perspectives about

self- and third-party governed platforms they use in their organisations; and third, we highlight the pivotal role that regulations (legislation and media self-regulation) play in news workers' decision-making processes.

##### 4.1. Audience Factor and its Related Practices

News media workers' emphasis on audience data varied among organisations and individuals. This highlights the ambivalent attitudes toward data practices in news media. For some, audience data significantly impacted their decision-making processes, which the following excerpt illustrates:

We base decisions less and less on emotions and feelings....Whether it's about planning journalism—what kinds of stories we want to make—or business decisions or designing our website—what kinds of functions we want to place there. In both cases, data is used diligently. (NW5)

However, not all interviewees put as much emphasis on audience data. They felt they had a responsibility to society to tell the most important news, no matter what metrics indicated, i.e., journalistic values guided their decision-making regarding the news. Audience data could guide the form of news stories, but newsworthiness and the story's content were news workers' decisions. In contradiction, some understood audience data as a means to emphasise particular kinds of content. For example, one interviewee said that they closely follow their audiences' actions by age group: "We follow what kind of content interests different aged audiences so that we can produce better content and are able to offer our subscribers the content they wish for" (NW9).

Interviewees' perceptions of audience data were indeed ambivalent: They were uncertain about how much data should and does affect their decisions and actions. Thoughts and perceptions about the influence of audience data also varied throughout the interviews, as news workers seemed to be weighing what they could say about data. Illustratively, one interviewee raised critical views of their sites' audience data use and presumed that it was unreliable regarding age, gender, and reading time statistics: "Reading time analytics is so unreliable that we can't base any decisions on it" (NW8). Later the same interviewee described that they could still partly rely on their (audience) data in cases such as planning the front page. They further described that analytics was understood as helping them to make decisions regarding their publication format or content:

We've got new kinds of formats such as live broadcasting and radio shows, and as the number of staff stays the same, we need to let go of something. This is how long-term analytics can help us decide which [formats/content] are not so important for us to do. (NW8)

Data-driven news culture was also visible in organisations' ways of testing news headlines and how different kinds of headlines allured their audiences. Some had even predicted, for example, how particular headlines performed for different aged audiences.

The origins of the data also mattered. Most importantly, the data of social media audiences raised concerns. Some saw a difference between the use of social media data and data that was collected from the organisation's sites—audience data guides more social media activities than activities on their own sites. They described this as a way to keep journalistic decision-making in their hands. In addition, audience data was utilised to compare different social media platforms' performance. As one participant observed, "It's interesting to compare how the same video performs and interests on TikTok and Instagram and what kinds of audiences it allures" (NW4).

Also, audience data is considered when practices relating to paywalls are decided upon. News media aims to increasingly understand their audiences' movements on platforms and ways to monetise them. In our study, interviewees from all three organisations said they employ hard paywalls. While the previous literature considers a "hard paywall" to be "no access to content without subscription" (Myllylahti, 2014) or that they "allow no free content" (Pickard & Williams, 2014), our interviewees seemed to connect hard paywalls to granting access to content that is only available to subscribers. Thus, they distinguished between some content that is freely accessible to all and clear-cut content that is unreachable unless readers subscribe. With that in mind, in this study, we use the concept of a "hard paywall," as our interviewees presented it. Further, in our study, the number of news stories placed behind the paywall varied between organisations. The first newspaper distributes a large amount of free content as they see access to accurate information as their journalistic responsibility in building a democratic society. The two other newspapers had a different approach to paywall practices, and most of their content was behind a paywall. They further described how the content is usually distributed for free in situations when the news originates from the Finnish News Agency (STT) or deals with global, national, or local security:

All the stories that we produce are primarily behind the hard paywall no matter where one enters the story, from social media or elsewhere....For free of charge, we offer things such as our columns, STT news, and for example, national instructions during Covid-19. (NW8)

The above-mentioned factors become apparent through newsroom practices—such as the paywall example illustrates—and manifest the central role of audience data in news workers' decision-making processes. The interviewees' statements reflect the contradictory

views of audience data's role in news work. The ambivalence and uncertainty toward audience data were also visible inside single organisations: Interviewees from the same news organisation shared strikingly different views on the impact of audience data in their decision-making processes. In addition, the statements portray how audience data is a post-publication factor that shapes new workers' decision-making which makes it an issue of (post-publication) gatekeeping—what kind of content news workers publish or should publish, and how they frame their journalistic content.

#### 4.2. Platform Factor and Its Related Practices

News organisations use a range of platforms to reach their audiences and distribute their content. Each platform has its unique mechanisms, and many are guided by algorithmic recommendations. When platforms were looked at through the lens of (post-publication) gatekeeping, the question of governance became evident—how much decision-making power the platforms afford to news workers. The interviewees described how organisations employed three different kinds of platform systems: self-governed, third-party governed, or a mix of these. On a self-governed platform (e.g., company websites and applications), the news organisations are fully in charge of the published content (what and when) and in control of dataflows (collection, storage, and handling; e.g., self-made analytics tool). From the viewpoint of paraphernalia, the materiality of technological objects is embedded in the platforms' software. That is, power over platforms is tied to the software being used: If a news organisation owns the software, it has more power to manage the platform it operates on. An effective way to gain control over data is to develop one's own analytics tools for the news organisation, as one interviewee states:

It's a platform we've developed for GDPR reasons. Because if we used another [third-party] platform, the data would go just somewhere....We've protected [the data] so that people can reply anonymously, and those answers are not connected to any other information within the organisation. (NW3)

Self-governed platforms were understood as means to control data, but third-party governed platforms were seen as governed by the technology giants. This means that news media have little power over the affordances that guide the use of the latter ones. The most often mentioned platforms were Google and Meta and their products: Google Analytics for following newsroom metrics; Facebook, Instagram, Twitter, and Tiktok for social media activities. Interviewees from each organisation described how much traffic in news organisations' sites originates from social media platforms. They were unified in that the main function of social media is to promote journalistic content, i.e., making their brand known.

However, understandings of platform governance varied. Some interviewees stated that they are not particularly utilising social media (data) since the organisation prioritises its own products, whereas some organisations put much effort into social media publishing. News workers also pay attention to the vernaculars of the platform. As one interviewee states: “We bring it [the news story] in a format that is easy to understand, e.g., to Instagram. So that it also supports visual storytelling. And that’s also shareable” (NW4). This highlights how the platforms’ affordances guide journalistic decisions. Interestingly though, some interviewees also highlighted social media as a tool for building a democratic society: “We fulfil our democratic duty by reaching young readers and by sharing trustworthy information with them about the regional elections. Social media has, in this way, a pure journalistic and noble purpose to serve democracy” (NW1).

As governance of platforms proved important, it also raised concerns about data reliability. This was particularly evident in the context of social media platforms. In some situations, social media companies’ data (e.g., Meta’s demographics) were deemed more trustworthy than the data of news organisations’ platforms, as this type of data was harder to collect from their own sites in a trustworthy manner. As one interviewee explains: “I trust age distribution [data] more on social media because I know that knowing their users is their [platforms’] main business” (NW8). Further, the reliance between social media platforms varied as one interviewee explains: “Google Analytics offers its own view about the visitors, but when you compare that to social media [Meta’s user data], it’s good to remember that the [Google] data only gives some ideas” (NW7). Data reliability was also deemed important when news organisations use platforms that are a mix of self—and third-party governed platforms. For example, news organisations have built their tools on top of Google’s infrastructure. As one interviewee explains:

Within our conglomerate, we’ve previously used a company which has built us analytics [tools]....They have utilised the data that comes through Google Analytics....But now we use an external company that tries to identify problems related to data reliability....They have found errors and problems in data collection and are now helping us fix them. (NW8)

The role of the platform factor becomes even more evident in news organisations’ moderation practices. News organisations seem to consider what the platforms afford them to do, i.e., do news workers need to moderate comments or does the platform or another service provider do it for them? The level of content moderation differed between newspapers. In the first news organisation, moderation for their website and Meta’s social media platforms was bought as an outside service and was conducted by humans and machine learning

software. Occasionally news workers moderate by themselves, e.g., on TikTok. News workers from the second news organisation said they had outsourced their website moderation, but they moderate their social media comments as part of their daily practices. News workers from the third news organisation explained that they do not currently use website commenting and that they rely solely on social media commenting in audience interaction. Social media moderation was seen as part of their daily practices: “Journalists who work online are in charge of social media moderation. You go there systematically to check the comments, but because you do that besides all other tasks, you can’t fully concentrate on it” (NW8).

The above-mentioned practices demonstrate that news workers base their decisions on platform data and vernaculars to some degree. Platforms can thus be seen as a post-publication gatekeeping factor that is a part of their decision-making process when deciding what to publish. Further, from the interviewees, it became evident that there is a hierarchy of trust towards different platforms. Organisations (and individuals inside them) value them differently based on the trustworthiness of the particular platform and its data practices.

#### 4.3. Regulatory Factor and Its Related Practices

As our theory-based extension to Hermida’s (2020) post-publication gatekeeping framework suggests, regulatory factors are part of the decision-making process of news workers, i.e., post-publication gatekeeping factors. The interviewees described how GDPR has extensively shaped their work practices. However, they gave considerably less attention to the aspects of journalism ethics and media self-regulation, which can also be regarded as part of post-publication gatekeeping. In the interviews, we asked direct questions, such as “how ethical and responsible do you think you are in your work.” In these situations, none of the interviewees mentioned the journalistic guidelines, the ethical codebook for mass media set by the Finnish press council, but rather discussed the effects of GDPR on their work. This is notable because media self-regulation often takes place *after* the news is published, for example, when a story needs revising based on feedback from sources. Hence, the press council also has a pivotal role in the media self-regulation process as it oversees how news media follow the journalistic guidelines *after* a news item has been published. The interviewees’ focus on GDPR, on the other hand, suggests that the law is an effective regulatory tool and part of news workers’ everyday work. This is illustrated by the following quote from an organisation that received feedback concerning the use of Google Forms from their audiences: “Yes, I feel we’re [ethical]. Earlier, we might have collected something through Google Forms but then gave up on it as it raises questions about data protection....[The] GDPR becoming effective was a turning point here” (NW2).

The quote also demonstrates that some audience members were concerned about regulatory issues. Furthermore, interviewees described how following GDPR has affected their working routines and increased audiences' rights over their data. This can be interpreted as a means for audiences to become part of the gatekeeping process. GDPR allows them to withhold the release of their personal data even after it has been handed over to news organisations and used, for example, in audience metrics. However, not all audience members seem to be equally aware of their rights, as one participant explains:

Users can ask us to empty all their data from our records...These requests might come once or twice a week. They are usually people who work in the IT field, experts that are interested in data. An average user isn't usually so interested in their data as long as it's in somewhat good hands. (NW5)

While interviewees brought up the theme of GDPR and how it has shaped news media work, they also described that GDPR and related legal issues were not their areas of responsibility. They repeated that GDPR was something that they did not need to be responsible for because their organisation has specific professionals for GDPR and legal matters. As one participant explains: "I'm not sure how to answer except that GDPR has brought along strictness...These sorts of issues are not our team's responsibility. I don't work with these issues" (NW2). This is noteworthy given that our interviewees were the ones that are dealing with data-related issues in their daily work.

Regulation's shaping of practices can further be exemplified through cookies that also shape the consumption of news after its publication. Due to GDPR, audience members can refuse to share their cookie data with news organisations and third parties who have made agreements with the organisations. Some interviewees believed that data will be regulated increasingly in the future and that regulation will bring changes to data use, both to social media companies and to news organisations themselves. They were worried about the demise of third-party cookies. This would make it difficult to access audience data and would thus challenge media organisations' business logic: "Google will close its support to third-party cookies, which means operating them will become more difficult...Also, in Finland, they [data protection officers] regulate the way we and other publishers operate now and in the future" (NW7).

At the same time, interviewees were after stricter regulation for the major platforms such as Google and Meta. They called for compensation for news content they had created, which now circulates for free on social media platforms in the post-publication context: "Many working in the industry hope that Google and Facebook would share their revenue with us, just as has happened in Australia and France in the past" (NW7). Interviewees hoped that future regulation would help their businesses, for example, news media could have

access to cross-platform data to see how their audiences migrate across their own and third-party platforms. At the moment, due to regulation and social media platforms' unwillingness to share their user data, news media organisations base their decisions on a single social media platform or their own platforms' data flows. This further demonstrates the gatekeeping power that these social media giants and regulations possess over news organisations.

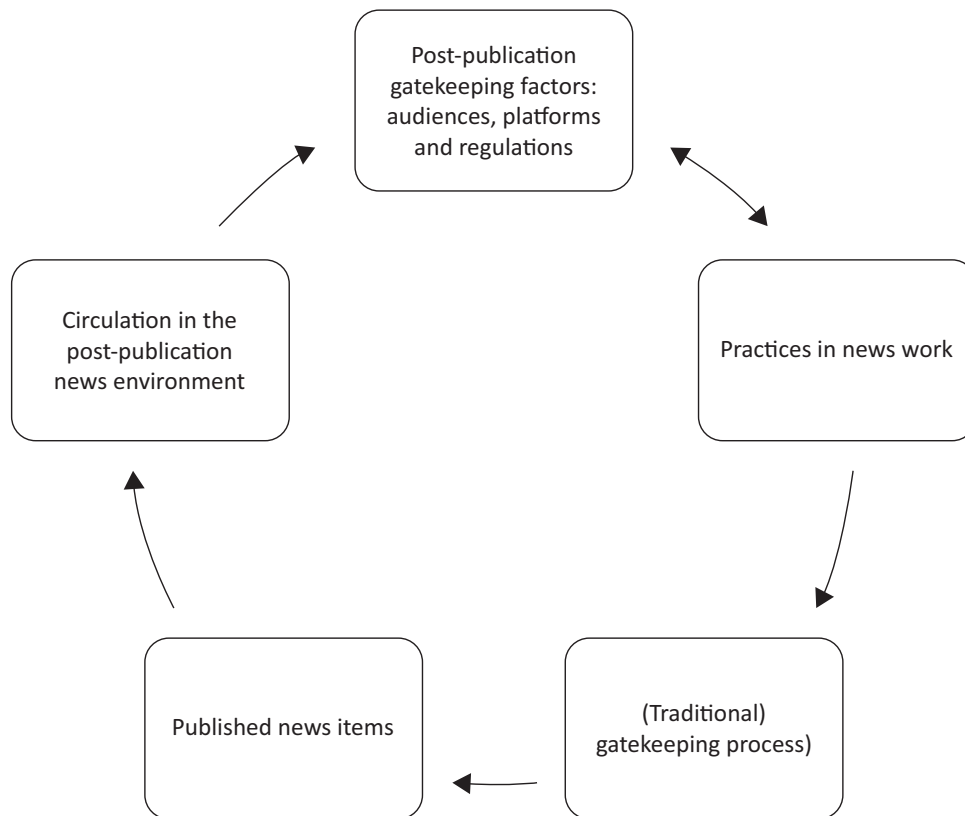
## 5. Discussion and Conclusion

In the previous section, we have demonstrated how the post-publication factors of audiences, platforms, and regulations, as well as their intertwined practices, iteratively shape news workers' editorial decision-making processes in datafied news work. We have highlighted the role of regulatory factors in the current (post-publication) gatekeeping processes and extended the framework of "Post-publication gatekeeping" by Hermida (2020) to include regulatory factors and practices in datafied news work. Our empirical findings support previous studies (e.g., Hermida, 2020; Salonen et al., 2022; Seuri & Ikäheimo, 2022) claiming that gatekeeping power is shared between multiple (f)actors in the current datafied news environment. To illustrate how the factors and practices are situated in the traditional journalistic gatekeeping process, we present Figure 2.

The figure illustrates how post-publication gatekeeping factors and practices iteratively shape the traditional journalistic gatekeeping process, which includes all the culling and crafting of information that takes place in newsrooms before a news item is published. After the news item has been published, it circulates in the datafied news environment. From there on, the news item interacts with audiences, platforms, and regulatory gatekeeping factors. These create new practices and shape traditional journalistic gatekeeping in news organisations. For example, the audience factor is present in the ways audience data guide decision-making over news headlines. The platform factor is present, for example, in cases where social media metrics guide what kind of content is published on a particular platform. Further, regulatory factors outline how audience data can be utilised. For example, audiences have the right to withdraw their data from news organisations' databases. These kinds of withdrawals shape audience metrics and can, therefore, also shape decision-making processes. For future studies, we recommend that researchers apply the suggested framework empirically to validate it further and more extensively map the factors and practices that shape post-publication gatekeeping processes in the datafied news environment.

The current study further contributes theoretically to the field of journalistic gatekeeping studies—it brings together the previous discussions of gatekeeping factors (e.g., Salonen et al., 2022; Wallace, 2018) and practices (e.g., Bruns, 2018; Singer, 2014) that shape the





**Figure 2.** (Post-publication) gatekeeping processes in datafied news work.

production and distribution of news. Furthermore, the study answers previous research's (Aral, 2020; Porlezza & Eberwein, 2022; Seuri et al., 2022) call to examine regulations in the datafied New Gatekeeping Regime. We have done this by not only suggesting the extension of regulatory factors to Hermida's (2020) framework but also by unravelling the news workers' contradicting perceptions concerning legal and ethical matters of their work. Our findings somewhat align with previous research (e.g., Ekström et al., 2022; Rydenfelt et al., 2022) on how news workers work with data; its use is negotiated with other, often journalistically-driven values. However, our study participants highlighted aspects of the law (i.e., the GDPR) more than media self-regulation, even when asked about the ethics of news work. It is noteworthy that while the Finnish press council is highly valued among news workers in Finland, it was not discussed in the interviews. This could be due to uncertainty arising from datafied news work. Several of our participants stated that they were unsure how data is used and that other non-editorial personnel might be better equipped to discuss legal aspects of their work. Data and knowledge were understood as being "out there."

Further, part of this confusion between law and ethics could be explained by the idea that the law (GDPR) limits the misuse of data collection and use and, therefore, forces more ethical approaches. Ethics is written into the GDPR, at least at the minimum. As the EU's new set of Digital Acts comes into force, the European

media market will face changes in business practices and models (Newman et al., 2022). These changes may also impact the future of media self-regulation. It is, therefore, crucial for news organisations to prepare for these changes and consider ethical aspects when planning their future. The interconnected nature of legislation and media self-regulation should also be the focus of further research.

Furthermore, our study confirmed that web analytics and the knowledge generated through audience data shape the journalistic gatekeeping process (cf. Tandoc, 2014), evident in practices such as personalising headlines and decisions relating to the front page. This supports the idea that journalism is shifting from an audience-centric view to a data-driven one, i.e., journalism (research) is experiencing a *data turn*. This is also connected to reliance on data and platforms. In our empirical data, the significant role of data processing technologies and, most notably, the various third-party actors who provide these technologies sheds light on what the digital trinity of algorithmisation, datafication, and platformisation (Latzner, 2021) looks like in everyday news work and how it informs the understanding of not only the audiences but also that of news workers concerning newsworthiness. This was visible in our findings, for example, as reliance on social media data.

However, reliance on data raises some critical concerns. First, our focus on news media organisations and their everyday working practices shows that the

understanding of the audience generated through data does not always reflect reality. This mirrors recent critical data and media research arguments on how knowledge created through data is, in fact, often ambiguous (Pink et al., 2018). Consequently, datafication may produce a fundamental misfit between data-generated understanding of an individual and their own experience (Talvitie-Lamberg et al., 2022)—to which the data double concept also refers (see e.g., Ruckenstein, 2014). Second, and as our study demonstrated, there is an ambivalence in how the news workers spoke of their data use, even inside a single organisation. This illustrates the controversial role of audience data in news work. News workers aim to base their decisions on platform metrics, but the knowledge of audiences is highly dependent on the (single) platform-generated data and the particular data processing tools (Aral, 2020). This means that journalistic decision-making becomes increasingly dependent on the platforms and their data processing practices. Further, this could lead to an undesirable situation where third parties are given too much power over journalistic decision-making (Salonen et al., 2022).

The knowledge of how data is used and exploited across different units in news organisations is fragmented. With further overall discussion on data-related practices, news organisations could form a less ambivalent relationship with the data they possess. As our results indicate, a move in this direction seems salient, given that audience data increasingly guides decision-making and newsworthiness in the news media. Therefore, involving news workers from different units and positions in discussions on data use and its regulatory aspects would help create a more holistic understanding of data and give individual news workers confidence in managing their responsibilities and expertise. We have begun to unravel this topic but also recognise the limitation of the size and nature of our dataset. Therefore, we invite future scholars to dig deeper into the ambivalence of data use in news workers' daily practices.

### Acknowledgments

The authors would like to thank the Media Industry Foundation of Finland, which funded the research, and the editors and reviewers for their time and comments that helped to improve the manuscript.

### Conflict of Interests

The authors declare no conflict of interest.

### References

Abbott, E. A., & Brassfield, L. T. (1989). Comparing decisions on releases by TV and newspaper gatekeepers. *Journalism Quarterly*, 66(4), 853–856. <https://doi.org/10.1177/107769908906600411>

- Andrejevic, M. (2014). Big data, big questions: The big data divide. *International Journal of Communication*, 8, 1673–1689. <https://ijoc.org/index.php/ijoc/article/view/2161/1163>
- Aral, S. (2020). *The hype machine: How social media disrupts our elections, our economy and our health—And how we must adapt*. HarperCollins.
- Berkowitz, D. (1990). Refining the gatekeeping metaphor for local television news. *Journal of Broadcasting & Electronic Media*, 34(1), 55–68. <https://doi.org/10.1080/08838159009386725>
- Breiter, A., & Hepp, A. (2018). The complexity of datafication: Putting digital traces in context. In A. Hepp, A. Breiter, & U. Hasebrink (Eds.), *Communicative figurations: Transforming communications in times of deep mediatization* (pp. 387–405). Palgrave Macmillan.
- Bruns, A. (2018). *Gatewatching and news curation: Journalism, social media, and the public sphere*. Peter Lang.
- Burgess, M. (2022, August 9). Will Europe force a Facebook blackout? *Wired*. <https://www.wired.co.uk/article/facebook-eu-us-data-transfers>
- Couldry, N., & Yu, J. (2018). Deconstructing datafication's brave new world. *New Media & Society*, 20(12), 4473–4491. <https://doi.org/10.1177/1461444818775968>
- Eberwein, T., Fengler, S., & Karmasin, M. (2019). Theory and practice of media accountability in Europe: An introductory overview. In T. Eberwein, S. Fengler, & M. Karmasin (Eds.), *Media accountability in the era of post-truth politics: European challenges and perspectives* (pp. 3–18). Routledge.
- Ekström, M., Ramsälv, A., & Westlund, O. (2022). Data-driven news work culture: Reconciling tensions in epistemic values and practices of news journalism. *Journalism*, 23(4), 755–772. <https://doi.org/10.1177/14648849211052419>
- Hanusch, F. (2017). Web analytics and the functional differentiation of journalism cultures: Individual, organizational and platform-specific influences on news-work. *Information, Communication & Society*, 20(10), 1571–1586. <https://doi.org/10.1080/1369118X.2016.1241294>
- Hermida, A. (2020). Post-publication gatekeeping: The interplay of publics, platforms, paraphernalia, and practices in the circulation of news. *Journalism & Mass Communication Quarterly*, 97(2), 469–491. <https://doi.org/10.1177/1077699020911882>
- Latzer, M. (2021). *Digital trinity—Controllable evolution—Everyday religion: Characteristics of the socio-technical transformation of digitalization*. Department of Communication and Media Research. [https://www.zora.uzh.ch/id/eprint/204675/1/Latzer\\_Digital\\_trinity\\_controllable\\_evolution\\_everyday\\_religion.pdf](https://www.zora.uzh.ch/id/eprint/204675/1/Latzer_Digital_trinity_controllable_evolution_everyday_religion.pdf)
- Meier, W. A., & Trappel, J. (2022). Media transparency: Comparing how leading news media balance the

- need for transparency with professional ethics. In J. Trappel & T. Tomaz (Eds.), *Success and failure in news media performance: Comparative analysis in the media for democracy monitor 2021* (pp. 255–273). Nordicom.
- Meraz, S., & Papacharissi, Z. (2013). Networked gatekeeping and networked framing on #Egypt. *The International Journal of Press/Politics*, 18(2), 138–166. <https://doi.org/10.1177/1940161212474472>
- Myllylahti, M. (2014). Newspaper paywalls—The hype and the reality: A study of how paid news content impacts on media corporation revenues. *Digital Journalism*, 2(2), 179–194. <https://doi.org/10.1080/21670811.2013.813214>
- Newman, N., Fletcher, R., Robertson, C., Eddy, K., & Nielsen, R. (2022). *Reuters Institute digital news report 2022*. Reuters Institute. <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2022>
- Ohlsson, J., & Facht, U. (2017). *Ad wars: Digital challenges for ad-financed news media in the Nordic countries*. Nordicom. <http://hdl.handle.net/2077/59562>
- Pantti, M. (2015). Visual gatekeeping in the era of networked images: A cross-cultural comparison of the Syrian conflict. In T. P. Vos & F. Heinderyckx (Eds.), *Gatekeeping in transition* (pp. 203–223). Routledge.
- Pickard, V., & Williams, A. T. (2014). Salvation or folly? The promises and perils of digital paywalls. *Digital Journalism*, 2(2), 195–213. <https://doi.org/10.1080/21670811.2013.865967>
- Pink, S., Lanzeni, D., & Horst, H. (2018). Data anxieties: Finding trust in everyday digital mess. *Big Data & Society*, 5(1), 1–14. <https://doi.org/10.1177/2053951718756685>
- Porlezza, C., & Eberwein, T. (2022). Uncharted territory: Datafication as a challenge for journalism ethics. In M. Karmasin, S. Diehl, & I. Koinig (Eds.), *Media and change management* (pp. 343–361). Springer.
- Roosa, S., Ritzer, C., Martin, N., Jansen, J., Rosenzweig, D., Schuitema, N., Filkina, N., & Ghebali, B. (2022). *European rulings on the use of Google Analytics and how it may affect your business*. Norton Rose Fulbright. <https://www.dataprotectionreport.com/2022/02/european-rulings-on-the-use-of-google-analytics-and-how-it-may-affect-your-business>
- Ruckenstein, M. (2014). Visualized and interacted life: Personal analytics and engagements with data doubles. *Societies*, 4(1), 68–84. <https://doi.org/10.3390/soc4010068>
- Rydenfelt, H., Haapanen, L., Haapoja, J., & Lehtiniemi, T. (2022). Personalisation in journalism: Ethical insights and blindspots in Finnish legacy media. *Journalism*. Advance online publication. <https://doi.org/10.1177/14648849221138424>
- Salonen, M., Olbertz-Siitonen, M., Uskali, T., & Laaksonen, S. M. (2022). Conversational gatekeeping—Social interactional practices of post-publication gatekeeping on newspapers’ Facebook pages. *Journalism Practice*. Advance online publication. <https://doi.org/10.1080/17512786.2022.2034520>
- Seuri, O., & Ikäheimo, H. P. (2022). *Gatekeeping in the digital age*. Sitra. <https://www.sitra.fi/en/publications/gatekeeping-in-the-digital-age>
- Seuri, O., Ikäheimo, H. P., & Huhtamäki, J. (2022). What happens when platforms mediate the audience–journalism relationship? In V. Manninen, M. Niemi, & A. Ridge-Newman (Eds.), *Futures of journalism* (pp. 227–243). Palgrave Macmillan.
- Shoemaker, P. J., & Vos, T. (2009). *Gatekeeping theory*. Routledge.
- Singer, J. B. (2014). User-generated visibility: Secondary gatekeeping in a shared media space. *New Media & Society*, 16(1), 55–73. <https://doi.org/10.1177/1461444813477833>
- Steedman, R., Kennedy, H., & Jones, R. (2020). Complex ecologies of trust in data practices and data-driven systems. *Information, Communication & Society*, 23(6), 817–832. <https://doi.org/10.1080/1369118X.2020.1748090>
- Talvitie-Lamberg, K., Lehtinen, V., & Valtonen, S. (2022). Tactics of invisibility: How people in vulnerable positions make datafied everyday life livable. *New Media & Society*. Advance online publication. <https://doi.org/10.1177/14614448221136077>
- Tandoc, E. C., Jr. (2014). Journalism is twerking? How web analytics is changing the process of gatekeeping. *New Media & Society*, 16(4), 559–575. <https://doi.org/10.1177/1461444814530541>
- Tandoc, E. C., Jr., & Vos, T. P. (2016). The journalist is marketing the news: Social media in the gatekeeping process. *Journalism Practice*, 10(8), 950–966. <https://doi.org/10.1080/17512786.2015.1087811>
- Tracy, S. J. (2018). A phronetic iterative approach to data analysis in qualitative research. *Journal of Qualitative Research*, 19(2), 61–76. <https://doi.org/10.2284/qr.2018.19.2.61>
- van Dijck, J., Poell, T., & De Waal, M. (2018). *The platform society: Public values in a connective world*. Oxford University Press.
- Vos, T. P. (2019). Journalists as gatekeepers. In K. Wahl-Jorgensen & T. Hanitzsch (Eds.), *The handbook of journalism studies* (pp. 90–104). Routledge.
- Vu, H. T. (2014). The online audience as gatekeeper: The influence of reader metrics on news editorial selection. *Journalism*, 15(8), 1094–1110. <https://doi.org/10.1177/1464884913504259>
- Wallace, J. (2018). Modelling contemporary gatekeeping: The rise of individuals, algorithms and platforms in digital news dissemination. *Digital Journalism*, 6(3), 274–293. <https://doi.org/10.1080/21670811.2017.1343648>
- Welbers, K., & Opgenhaffen, M. (2018). Social media gatekeeping: An analysis of the gatekeeping influence of newspapers’ public Facebook pages. *New Media & Society*, 20(12), 4728–4747. <https://doi.org/10.1177/1461444818784302>

## About the Authors



**Margareta Salonen** (MA) is a doctoral researcher in journalism at the Department of Language and Communication Studies, University of Jyväskylä, Finland. Her work focuses on the triadic power relations between news media, audiences, and social media platforms. More precisely, she researches social interaction between news media and audiences, datafication of news work, and journalism ethics on social media platforms. She is a member of the Finnish press council and the Student and Early Career Representative of the ICA Journalism Studies Division.



**Veera Ehrlén** (PhD) is a postdoctoral researcher in journalism at the Department of Language and Communication Studies, University of Jyväskylä, Finland, where she also earned her doctorate in journalism in 2021. She has expertise in studying mediatisation of leisure cultures, digitalisation of social interaction, social media consumption and content distribution practices, and news consumption in data-driven digital environments.



**Minna Koivula** (MA) is a doctoral researcher in journalism at the Department of Language and Communication Studies, University of Jyväskylä, Finland. Her research examines journalism through the lens of organisational communication research and explores the shifting institutional context of work in the media field.



**Karoliina Talvitie-Lamberg** (PhD) is an assistant professor in media, communication and journalism at the Department of Language and Communication Studies, University of Jyväskylä, Finland. Her research interests include everyday datafication in media publics, inclusion and vulnerabilities, data ethics, digital disconnection, and AI and bot-enhanced communication. She is a specialist in digital inclusion issues in media and public service development. Her work has been published in several esteemed journals, including *New Media and Society*, *Computers in Human Behavior*, and *Surveillance and Society*.