

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Ahva, Laura; Salonen, Margareta; Ovaska, Liisa; Talvitie-Lamberg, Karoliina

Title: The managerial and future-oriented role of audience data in data-informed news organisations

Year: 2024

Version: Published version

Copyright: © The Author(s) 2024

Rights: CC BY 4.0

Rights url: <https://creativecommons.org/licenses/by/4.0/>

Please cite the original version:

Ahva, L., Salonen, M., Ovaska, L., & Talvitie-Lamberg, K. (2024). The managerial and future-oriented role of audience data in data-informed news organisations. *Journalism*, Early online. <https://doi.org/10.1177/14648849241285721>

The managerial and future-oriented role of audience data in data-informed news organisations

Journalism
2024, Vol. 0(0) 1–19
© The Author(s) 2024



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/14648849241285721
journals.sagepub.com/home/jou



Laura Ahva 

Tampere University, Finland

Margareta Salonen 

University of Jyväskylä, Finland

Liisa Ovaska 

Tampere University, Finland

Karoliina Talvitie-Lamberg 

University of Jyväskylä, Finland

Abstract

One sign of datafication in journalism is that newsrooms utilise insights from audience data and related metrics in their decision-making. In professional discourse, this often alludes to the notion of ‘data-informed newsrooms’ in which organisations are managed with the help of data. This article focuses on the managerial role of audience data in such organisations: What types of audience data inform newsrooms, which domains of newsroom management are being informed by data, and what kinds of visions motivate the future of data use in news organisations? These questions are answered based on empirical material from three different research projects (2020–2022) in Finland, including both qualitative interviews and survey results featuring the views of varied journalism professionals on data use. The theoretical framework draws from conceptualisations of anticipatory governance. Professionals’ interpretations of audience data reveal the basic elements of data sources, types, systems and storages that inform the organisations. Such audience data are used for four main management purposes—leading individuals, work processes, the news product, and the audience. The role of audience

Corresponding author:

Laura Ahva, Faculty of Information Technology and Communication Sciences, Kalevantie 4, Tampere University, Tampere 33014, Finland.

Email: laura.ahva@tuni.fi

data in the future preferably includes elements of predictability, fairness, usability, and depth. However, there are discrepancies between the visions themselves as well as the views of the managers and journalists, which results in an ongoing negotiation about the suitable role of audience data in news organisations.

Keywords

Anticipatory governance, audience analytics, data-informed newsroom, data-driven management, digital journalism, newsroom management

Introduction

Collecting and analysing digital data from all walks of life have become easier and much more effective in recent decades. This shift, often termed *datafication* (Couldry and Hepp, 2017), is also shaping the production and consumption of news and journalism (Porlezza, 2023). Newsrooms are, thus, among the many organisations aspiring to become ‘data-driven’ (Anderson, 2015). This refers to how their operations prioritise and give status to values and practices that derive from data and their analysis (Ekström et al., 2022: 757). Digital data are perceived as a central knowledge resource that needs to be harnessed, processed, and structured so that organisational actors in newsrooms can act upon them (Vulpius, 2022: 353). Digital data have an especially strong role in the work of newsroom managers who regularly utilise them to lead their organisations (Christin and Petre, 2020; Tenor, 2023). In such management, data and analytics help anticipate the future better (Flyverbom and Garsten, 2021).

It is evident that newsrooms have adopted insights from a broader *paradigm of data-driven management*. McAfee and Brynjolfsson (2012: 4) argued that access to big data brought ‘a revolution’ to the world of management at the start of the 2010s, suggesting that through big data, ‘we can measure and, therefore, manage more precisely than before’. Even if the revolutionary discourse has attenuated, a data-driven management paradigm remains in the information systems and business sectors. Based on the literature from this domain, data-driven organisations appear as entities that utilise data and analytics to increase their capabilities on four key organisational domains: structures, technologies, processes, and people (Berndtsson et al., 2018; Hupperz et al., 2021). The literature considers data a central asset, particularly in terms of predicting, rationalising, and optimising workflows, as well as building new products or finding new business models to improve productivity (Hannila et al., 2022). It has been posited that organisations need to be led in a way that cultivates a company-wide data culture to avoid silos and misunderstandings (Berndtsson et al., 2018).

We can identify two specificities in how the above-described paradigm has been domesticated into newsrooms. The first is *audience-centrism* (e.g. Lindén et al., 2022; Vulpius, 2023) and pertains to the fact that the key form of data in news organisations deals with users rather than any other type of organisational or business data (see Napoli and Roepnack, 2018). Audience data are appealing because they create the promise of

helping journalists gain control over their online audiences and, thus, better address the audiences' needs (Ahva and Ovaska, 2023; Anderson, 2011). However, we also know that the introduction of audience analytics into newsrooms has required news professionals to ponder the ensuing implications on their professional practices and values, such as autonomy (e.g. Christin and Petre, 2020). To summarise, journalists have weaved audience analytics as part of their profession through a process that started with a weak interest in the technology (MacGregor, 2007: 283), followed by resistance (Welbers et al., 2016) that slowly gave way for negotiation, interest, and eventually even enthusiasm (Ahva and Ovaska, 2023; Cherubini and Nielsen, 2016). As a result, analytics have become an everyday feature in newsrooms (Tenor, 2023).

However, the views about how forcefully data could and should drive news organisations' decision-making and management remain mixed (Ahva and Ovaska, 2023; Bunce, 2019; Salonen et al., 2023; Tandoc, 2019: 43; Vulpius, 2022). Hence, the second specificity deals with adjusting the notion of being 'driven' or 'led' by data into being 'informed' by data in the journalistic context. *The Guardian's* audience editor Chris Moran explicitly alluded to this in a report by Cherubini and Nielsen (2016: 32):

We describe ourselves as data-informed, not data-led, and that's critical to me. I'm obsessed with data in the newsroom, but it shouldn't be the only thing that is making the decisions: editorial instinct should do that.

Existing literature shows how journalistic work requires professional assessment of factors such as newsworthiness, importance, or impact, thereby featuring an editorial understanding of news as public information, which is difficult to assess using data analytics alone (Hollifield, 2019; Sirén-Heikel et al., 2022: 337–338). e.g., based on a study in a local Swedish newsroom, Ekström et al. (2022: 758) concluded that news organisations indeed embrace and learn from audience analytics and the metrics produced but only in combination with other values in making journalistic judgements. Similarly, Milojevic (2022) suggested that the aim of being data-informed has manifested in Norwegian newsrooms, for example, in how part of the front page is always reserved for stories that are chosen based on editorial assessment and not metrics. Lindén et al. (2022: 334) added that in many Nordic newsrooms, qualitative audience analyses are being combined with quantitative measurements to avoid being purely data-driven. So, it seems that news professionals feel ambivalence toward how much data should be allowed to guide their daily practices (Salonen et al., 2023).

Therefore, in newsroom discourse, the data-driven management paradigm has been reduced into a data-informed framework. In existing journalism research, the notion of 'data-informed newsroom' is referred to (e.g. Hendrickx et al., 2021), but not often unpacked. Cherubini and Nielsen (2016: 10) define it broadly as a journalistic organisation whose combination of technological tools, organisational structure, and newsroom culture enables data-based decision-making.

Newsrooms are thus currently giving audience data a significant informing role, and we need to get a better empirical and analytical hold of the core questions of what is it that data actually guide in data-informed newsrooms, and how is such informing justified.

Much is known about the role of audience data in journalistic work (e.g. [Christin, 2020](#); [Tandoc, 2019](#)), whereas the broader perspective of management has gained less attention (see however, [Bunce, 2019](#)). This article contributes to journalism studies by unravelling the components of data-informed newsrooms by shedding light on data's managerial role in the present, and visions of data for the future. Through this, we aim to understand what datafied newsroom culture means in practice. As our theoretical lens, we use literature on data-driven management, particularly the notion of anticipatory governance ([Flyverbom and Garsten, 2021](#)).

In the article, we focus on how journalism professionals make sense of the basic elements of audience data, and the domains of data-informed newsroom management, and how they weave data as part of the newsrooms' future in the cultural and societal context of Finland. The research questions are:

RQ1: What kinds of audience data are currently collected and used to manage news organisations in Finland?

RQ2: What domains in newsroom management are being informed by audience data?

RQ3: What kinds of future visions do the news professionals have about data-informed news organisations?

Data-informed management as a form of anticipatory governance

As we focus on the informing role of audience data in newsrooms, we are interested in the management rather than production practices of datafied journalism. Within the field of media management, we focus on a specific area of *newsroom management*: how the processes of creating news content are directed and coordinated ([Hollifield, 2019](#): 3). Our approach taps into two major themes in newsroom management studies: the role of technology and data in management, and the changes caused by the economic decline in the news industry ([Hollifield, 2019](#)). Currently, pressing issues in Finland pertain to managing newsrooms that focus more heavily on online subscriptions as a form of income ([Lindén et al., 2022](#)) as well as optimising user experience ([Leino et al., 2023](#)). Both issues are developed with the help of data analytics.

To examine these datafied changes, journalism researchers can learn from critical data scholars who have noted that when data, algorithms, and platforms become part of organisations and our everyday lives through datafication, *anticipation* appears as the central feature in three interlinked ways: in how people anticipate their actions, how they imagine the world, and especially how prediction through data has become a central way to do business and governance ([Katzenbach and Ulbricht, 2019](#); [Koivunen et al., 2024](#)). Critical data scholars posit that such transformations in the relationship between data and knowledge production ([Flensburg et al., 2023](#); [Kitchin, 2014](#)) and the power of data ([Andrejevic, 2019](#)) should be at the centre of research on datafication.

Anticipation is indeed a factor in the development and management of datafied journalism, too. [Napoli \(2016\)](#) argued that within journalistic organisations, data's role in

management can be seen as the latest step in how newsrooms have rationalised audience understanding in their need to operate among uncertainties. [Ananny and Finn \(2020: 1609\)](#) added that news publishers largely agree that the economics of online news require them to manage their organisations via anticipatory infrastructures, for example, to predict audiences' attention and engagement with metrics to better adjust the workflows of newsrooms.

To understand this anticipatory transformation in newsroom management, we can find further help from organisation and management research that indicates how the emergence of digital data as a central anticipatory asset shapes organisations: the way data constructs the organisations and carries them into the future ([Flyverbom and Garsten, 2021](#)). Data analytics and automated forms of pattern recognition have given rise to a new style of management in which organisations seek to make phenomena, problems, and opportunities visible, predictable, and, hence, manageable via data ([Flyverbom et al., 2016](#); [Flyverbom and Garsten, 2021](#)). Therefore, the means for managerial control and the exercise of power have also become data-informed. However, in this line of thinking, management takes place subtly via steering data-based processes or offering templates for organisational action, not only or necessarily via regulating individual employees' actions ([Flyverbom and Garsten, 2021: 2](#); [Flyverbom et al., 2016: 101](#)).

Following [Flyverbom and Garsten's \(2021\)](#) work, we posit that data are a central, albeit not the only, knowledge resource through which news organisations are governed and that the main aim of such governance is to anticipate, to lean forward. [Flyverbom and Garsten's \(2021: 7\)](#) work offers a four-level typology in which data-informed anticipatory governance creates a steering order in organisations. First, there are snapshots: statistical reports, rankings, visualisations, or other indications that help to simplify complexity and guide *immediate day-to-day actions*. Second, automated and predictive data analyses, such as profiling, provide prognoses to help predict *the near future*. Third, data are analysed to form projections and scenarios to actively design *the future ahead*. Fourth, data play a role in how *distant futures* are visioned, even in speculative and imaginary ways that often draw from external sources, such as data consultants' formulations (see also [Beer, 2019](#)). [Flyverbom and Garsten \(2021: 18\)](#) indicated that their typology can be applied to studying how data inform and shape organisational strategies, innovations, and the organisations themselves through sensemaking. We are particularly inspired by the last aspect.

In the context of data-informed newsrooms, the four levels of anticipatory governance appear as patterns that give structure to the managerial use of audience data. They invite us to regard news professionals' data-related interpretations as sensemaking that crafts these organisations in terms of anticipating both the immediate next moves and the near future, as well as more distant futures. We argue that sensemaking about anticipatory governance is at play when journalists and managers negotiate between journalism- and data-related values while aspiring to be data-informed rather than data-driven (see e.g. [Hendrickx et al., 2021](#); [Rolandsson et al., 2022](#); [Salonen et al., 2023](#); [Tandoc 2019](#)). Thus, journalists and newsroom managers are pushed to consider the appropriate place and weight of data in their current work and in planning for their organisations' near future (e.g. [Christin and Petre, 2020](#); [Chua and Westlund, 2019](#); [Hayes and O'Sullivan, 2023](#)). At the same time,

they also reflect on the many visionary promises of data and analytics (Ekström et al., 2022: 756; Svensson, 2023).

Research material and methods

The research material comprised three data sets from separate projects based in Finland between 2020 and 2022 (see Table 1). Despite some differences in the original research designs¹, all the projects had an overarching aim of understanding the use of audience data within Finnish news organisations, which gave us the idea of examining the data sets jointly. By focusing on Finland, we contribute to knowledge about newsroom management in the context of a Nordic journalism culture that used to be socially and economically stable, but in which newspapers in particular—even if still well-regarded by the audience both as print and online versions—have faced a decline in circulation (Hellman, 2021). In this context, newsrooms have actively utilised and developed data analytics (Milojevic, 2022; Tenor, 2023). Most commercial news organisations rely on two main sources of income: advertising sales and subscriptions, with the latter on the rise (Lindén et al., 2022). The proportion of people paying for online news is higher in Finland than the average in Nordic countries (Newman et al., 2022). Therefore, this study also sheds light on environments beyond the Nordics, in which subscription-based models are applied in online journalism.

Data Set 1 comprised nine semi-structured interviews with news managers (editors and unit leaders) and one supportive technologist (digital-domain specialist) working within Finnish news organisations. This data set gave us knowledge about the professionals responsible for buying services and/or implementing the use of analytics, that is, the

Table 1. Data Sets and informants of each research project.

Data sets		Data set 1 (N = 9)	Data set 2 (N = 9)	Data set 3 (N = 105)
Position	Managers (e.g. editors, department heads)	8	5	27
	Journalists (e.g. reporters, visual journalists)	0	1	75
	Supportive/technological staff (e.g. analysts, developers)	1	3	3
Gender	Female	2	1	51
	Male	7	7	50
	Nonbinary/prefer not to say	0	1	4
Type of news organisation represented	Newspaper	6	9	105
	Public broadcaster	1	0	0
	Tabloid	2	0	0
Time of interview/response		12/2020–2/2021	5/2022–8/2022	2/2021–11/2022
Average length of interview/response rate		69 min	96 min	17.5%

individuals who lead with data. This material was collected through remote interviews. The questions in the interview guide were organised under three main themes: (1) managers' and their newsrooms' data-related work practices, (2) audience metrics' role in Finnish news outlets more generally, and (3) the balance between in-house and outsourced data analysis within their companies.

Data Set 2 comprised nine semi-structured interviews with news professionals who dealt with data-related questions in their everyday work. Their work profiles varied from managerial roles (e.g. head of news) and content-producing journalists (e.g. social media manager) to supportive or technological roles (e.g. lead developer). Thus, this set offers knowledge about how experts in audience data within their units regard data-informed management. Some interviews were conducted via Zoom, and others were face-to-face. All were organised under six main themes: (1) collection and use of data from news media websites, (2) and social media platforms, (3) ethics and responsibility in data collection and use, (4) news media's approach to social media platforms, (5) legislation dealing with data practices, and (6) moderation of news discussions.

Data Set 3 comprised online survey responses from 105 news professionals with a 17.5% response rate²—including journalists, managers, and supportive staff—from one national newspaper and four regional newspapers. The survey included both Likert scales and open-ended questions covering (1) general habits following and utilising analytics, and (2) their impact on newsroom practice, (3) reasons to follow analytics, (4) emotions evoked by them, and (5) the information that analytics generate from the audience. This data set provided us with insights from a variety of news professionals, including those whose work is being managed.

Informed consent was obtained from all participants in each project separately³. All the qualitative interviews were transcribed and anonymised for analysis. The material was analysed so that a qualitative analytical framework was created by the first author based on Data Set 1 to cover interpretations of (1) sources and types, (2) present management domains and (3) future uses of audience data. Thereafter, the second and third authors followed the framework and conducted analyses of their qualitative material (interviews and open-ended responses). After the qualitative part, statistical distributions in the survey were used to confirm or challenge the categorisations. The authors then iteratively reinforced and nuanced the framework based on the individual data sets and created analytical categories that applied to all the research material.

Basic elements of audience data in newsrooms

Before delving into data-informed management, we address our first research question and clarify what news professionals viewed as valid audience data within their organisations and how it was handled. In [Figure 1](#), we condense these basic elements into main types of audience data sources, objects, systems, and storage methods originating from our respondents' interpretations.

The most important sources or sites of audience data comprised news organisations' own websites and mobile apps. However, data gathered from the organisations' purchase and subscription systems have grown in importance to specify the types of content

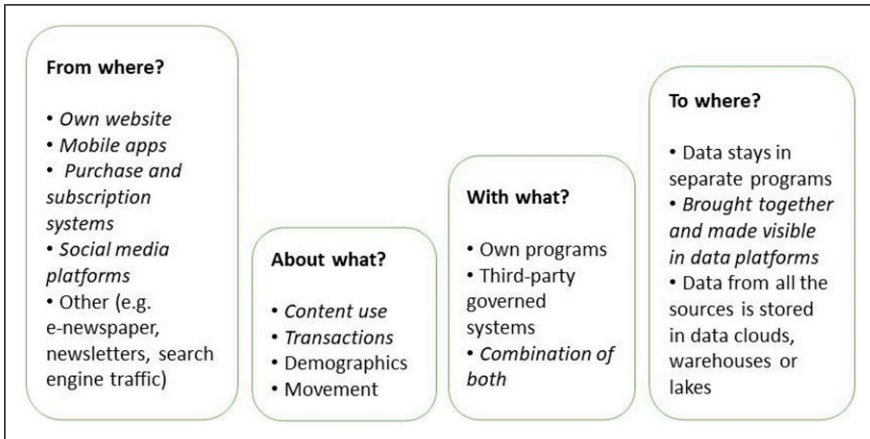


Figure 1. The main data sources, types, systems, and storages in Finnish newsrooms. The most typical varieties have been italicised.

subscribers consume or what would be converted into purchases of single stories or trial/full subscriptions. Social media platforms were also important data sources, although the relationship between news organisations and social media platforms feature dependencies and tensions (see [Chua and Westlund, 2019](#)).

In terms of types, content use data—page views, reading time, shares, visiting frequency, and sessions—were prominent, followed by transaction data, for example, purchases. There were also frequent references to demographic data about age, gender, residence, etc. However, as the collection of such data is increasingly regulated in the European Union (EU), this data type appeared to be the most difficult to obtain but also very valuable. Furthermore, movement data were collected to follow users' movements within news services or from social media to news sites.

Only large media companies had in-house analytics programs designed just for them; smaller organisations relied on services provided by externally governed third-party companies. In our material, about 20 different companies and programs were mentioned, the most frequent being Google Analytics, Chartbeat, LinkPulse, Tableau, Microsoft Power BI, and Adobe Analytics. However, the most typical case is a combination of internally and externally governed systems. In terms of storage, representatives from smaller newsrooms mentioned needing to operate with the help of separate analytics programs. However, most often organisations utilised data platforms, that is, software solutions that connect various programs and visualise the results (see [Flyverbom et al., 2016](#)). Lastly, only representatives from large media companies mentioned large-scale storage systems, such as data warehouses, clouds, or lakes.

Overall, [Figure 1](#) provides an overview of audience data in Finnish newsrooms. However, coherence and precision were not easy to find in the professionals' interpretations; hence, no order readily emerged from the analysis (see also [Milojevic, 2022: 7](#)). For example, although some emphasised the importance of specifying different

audience data sources (especially between social media and their own websites), these differences often vanished in the talk, with a typical reference only to ‘the audience’. This implies that most of the sensemaking about audiences was based on a category that refers to a mass understanding of *the online audience*, even if the wish towards specificity was clear in terms of increased value put on the categories of *registered*, and especially, *paying audiences*.

Data informs the management of individuals, processes, news products, and audiences

Our analysis revealed that the higher up a professional was within the organisational hierarchy, the more engaged they were with data and analytics. Compared with the group we refer to as managers (see [Table 1](#)), journalists used analytics less, and they had little or no authority to decide what types of data they received and which analytics software to use (see also [Hendrickx et al., 2021](#)). The decisions about which programs would be bought from third parties and which developed in-house, were made in negotiations between the managers and technologists, who also viewed the data-informed management approach as beneficial, while journalists were still coming to terms with it ([Christin and Petre, 2020](#); [Rolandsson et al., 2022](#)). Nevertheless, all groups shared an understanding of the main domains whose management was data-informed. To address our second research question, we suggest four overlapping but analytically separable domains that were regarded as informed by data in newsroom management.

The first domain consists of *individuals*. Here, the aim of leading with and through data is to guide journalists’ work and professional development as writers and providers of news. Thus, data are meant to inform journalists to become more skilled in news work. No examples of bonuses or control measures for individual journalists based on analytics were indicated in our material (cf. [Tandoc, 2019](#): 43). Instead, within many of the studied news organisations, management appeared ‘soft’ ([Petre, 2021](#): 170); that is, journalists individually followed the metrics and received a monthly list of their top-performing stories in terms of views and subscriptions, and individual coaching discussions could be arranged (see also [Ekström et al., 2022](#)). This highlights how data about the past turn into algorithmic knowledge that directs the immediate and near-future actions of individuals ([Flyverbom and Gartsen, 2021](#)).

Thus, insights from audience data were used to motivate and educate journalists to produce stories that aligned with metrics in a way that did not necessarily feel like authoritative management ([Flyverbom et al., 2016](#); see also [Hayes and O’Sullivan, 2023](#): 13; [Petre, 2021](#): 184). Furthermore, many participants expressed how they had become better journalists due to analytics, as they had learned to write headlines or story leads, develop story ideas, or choose angles that raise interest among audiences. Overall, about 70% of the survey respondents in Data Set 3 said they had made changes to their work based on analytics and that they had assessed their own success with the help of metrics. The practice of headline testing (AB testing) ([Ekström et al., 2022](#): 764) also fell into this category in situations where audience metrics indicated the preferred headline and thereafter guided the work of individual editors and managers. These professionals

educated and managed themselves to become better headline writers: *'[I use metrics] to learn how to make headlines that do well'* (Manager 12, Data Set 3).

The second domain deals with *work processes*. This is different from the previous category, in which data were used to improve individuals' skills. Here, the data inform the ways in which newsrooms' workflows are organised. This was apparent, for example, when the informants talked about how data guided (or should guide) the use of time. Simply put, work deemed successful by the metrics was allocated more resources (Petre, 2021: 124–125). Similarly, decisions to renounce certain practices were guided by insights from the data analysis, as the news editor in Data Set 2 exemplifies:

We can also think about how much time is used to make a story, and if we clearly see [from the audience data] that the kind of story format raises less and less interest, then we can decide if we should use that time for something better and more important. (News Editor 8, Data Set 2)

The publishing pace and timing were also fine-tuned according to the metrics to fit with the rhythms of audience visits on the site (Ekström et al., 2022: 764), a clear indication of anticipatory logic. Based on our material, this kind of work process management typically happened at the department or unit level. In the survey, almost all respondents agreed with the proposition that the success of a newspaper's various sections is assessed against audience data. Furthermore, one of the managers in Data Set 1 mentioned playful competitions between different departments to reach daily goals for new subscriptions. Here, the game-like following of transaction data aimed to motivate departments (cf. Individual 'traffic game'; Petre, 2021).

However, the most conspicuous element in this domain was how feedback practices have become datafied (see also Vulpius, 2022: 355). Teams or sections typically received listings of the best-performing articles on a weekly basis, and these listings were used as the basis for collective feedback. The surveyed news professionals expressed how most attention in such editorial feedback was paid either to the bunch of stories that did really well or to those that did poorly per metrics, and they pointed out—with irritation—how those that fell between these extremes were largely ignored (see also Bunce, 2019; Tandoc, 2019). However, using audience data to identify long-term trends in user behaviour to help plan newsroom work accordingly was viewed as beneficial. This illustrates how data-informed management operates by making trends visible to modify future actions (Flyverbom et al., 2016).

The third category refers to how the *news product* is managed. In this context, the news product appeared to comprise both the content and technical service of the site or app. One of the managers in Data Set 1 mentioned that the old saying 'Content is king' no longer applied because the combination of content *and* service had now become king. Both play a role in providing a pleasant user experience (Leino et al., 2023; Vulpius, 2023: 16), particularly one that entices the user to spend a large amount of time on the website (Hendrickx et al., 2021). Whereas technologists focused on managing technical service, most interpretations in our sample pertained to the content dimension of the news product.

Audience data from websites and social media platforms clearly informed decisions pertaining to news content, particularly in terms of topic selection or choosing stories that

would deserve follow-ups, as well as developing entirely new story concepts. Further, decisions regarding which stories should be placed behind the paywall were heavily datafied (Hendrickx et al., 2021). The practice of following metrics in real time (Vulpius, 2022: 355) to immediately tweak the presentation of already-published articles was also mentioned frequently, entailing the ordering of stories on the home page, wording of headlines, ‘refreshing’ of older articles’ leads, etc.:

By following metrics, we can react to bad figures, and we do this most typically by changing the headline or observing that certain topics just do not engage readers. (Manager 32, Data Set 3)

The fourth domain is the *audience*. The data gathered from the users were directed back at them with the goal of managing their future actions (cf. Flyverbom and Gartsen, 2021: 10–12). This was evident, for example, in how personalisation and recommendation systems were discussed. With prior knowledge about their content use on the service, audiences were offered certain types of stories, for example, by ordering the front page differently for different people (see also Svensson, 2023). However, this type of automated personalisation service was present only within the biggest news outlets.

Data also informed how the newsrooms conceived their audiences (Anderson, 2011). The introduction of subscription metrics seemed to have changed how various audience categories were managed. According to our analysis, users were guided by outlets to move from the general category of the online audience to that of the registered audience to gain more background information about them, but the (potentially) paying audience was catered to the most and in the most detailed fashion. According to a manager in Data Set 1, paywalled stories and related subscription metrics had brought along the practice of viewing paying audiences as smaller target segments or profiles. For example, stories were designed around ‘35-year-old mothers’ or ‘25–30-year-old women on Instagram’. Chosen topics were also sponsored on social media platforms for better targeting. Thus, news audiences were discussed and treated in much the same way as customers in marketing.

The future of data-informed newsrooms rests on predictability, fairness, usability, and depth

Regarding our third research question, we analysed the respondents’ visions and expectations about the managerial role of audience data that pertained to distant futures (Flyverbom and Garsten, 2021). This section is primarily based on Data Sets 1 and 2 because the survey structure of Data Set 3 did not provide space for imagining. However, the material often featured critical counterparts of the visions. Thus, the following scenarios are based on the analysis of the first two data sets, reflected against the dissatisfaction explicit in the third one.

The first vision concerns the *predictability* sought from the audience data. This echoes a way in which analytics companies view data as ‘prophetic’ (Beer, 2019). Predictive analytics refers to ways of analysing historical data with the help of data mining, statistical

modelling, or machine learning to guide future activities (e.g. [Berndtsson et al., 2018](#)). This vision is thus linked to the domains of managing individuals and news products. In the newsroom context, this refers to the anticipation that audience data could be used not only to assess how news outlets performed last week but also how they could foresee them performing better in the future ([Vulpius, 2022](#)).

In smaller newsrooms that utilised more rudimentary tools ([Cherubini and Nielsen, 2016](#)), predictability was still a wish or a fantasy, whereas bigger organisations were already assembling data to predict the future with some accuracy (cf. [Flyverbom and Garsten, 2021](#)). They developed and tested ‘editorial bots’ that suggested better wording, the most suitable story elements, or articles to be placed behind paywalls. However, such development work was still viewed as expensive and slow. Machine learning requires quality training data for reliable programs and, more importantly, the staff’s ability to envision the potential applications of such technologies for journalism.

The second vision pertains to the *fairness* of data collection and use and deals with management in three domains: audiences, work processes, and individuals. Regarding the last, the participants recognised the ‘data-driven’ risk of metrics becoming a managerial tool for disciplining individual journalists by measuring their performance based on numbers and not journalistic values (see [Petre, 2021](#)). Fairness was envisioned to counter this risk and to happen by making newsrooms’ basic data elements ([Figure 1](#)) clearer and more transparent to all, not just to managers who lead (see [Flyverbom et al., 2016](#)). In regard to how work processes were managed in different units, a journalist (Respondent 34, Data Set 3) noted that ‘*analytics is a tool for the mutual competition of the bosses*’, suggesting that journalists did not necessarily view the aforementioned subscription competitions as fair, motivational, or relevant management practices. The managers observed this to a degree and expressed the self-criticism that for a fruitful data-informed newsroom culture to emerge, they need to be more level-headed to avoid unnecessary hype around web analytics that can alienate journalists.

However, the fairness vision mostly pertained to the domain of managing audiences, as the informants emphasised that newsrooms should practise data collection based on the appropriateness principle: data should not be collected if it is not to be used. Privacy was also noted: Some participants expressed how personal details, such as age and gender, should not be collected by default. However, they were a minority, as the general ethos cherished the idea of knowing even *more* detailed information about target audiences. There were also visions of further developing the subscription-based business model to replace the current practice in which audiences pay for news with their personal data. One technologist emphasised that there should not be a price on private information:

There is no general measurement of how a person should get paid in exchange for a childhood memory, their date of birth, or the information that they like role-playing or whatever. There is no price for that. There are things we should not have to do business with. In my opinion, individuals’ private information is like that. (Data Specialist 5, Data Set 2)

Other interviewees also envisioned a situation in which news organisations could serve audiences more fairly by offering them back a set of analysed information about their

personal news use. For example, they referred to how the music-streaming platform Spotify annually compiles information about users' streaming and provides identity-building resources for listeners. Manager 1 in Data Set 1 envisioned an idea in which news organisations would offer audiences a service called 'Media Me', in which media companies could also take on a more public service-oriented and transparent role—like open data movements (Baack, 2015)—in giving back some of the data collected from users.

The third vision in the material pertains to *usability* and resonates with how analytics companies foster 'accessibility' as a key element in their data imaginaries (Beer, 2019). The news professionals fantasised about finding all the relevant data for journalistic work in a single location. Manager 2 (Data Set 1) expressed the worst scenario of having to follow 10 different systems separately. An ideal scenario would be an integrated service (see also Milojevic, 2022) that features a more fluent incorporation of social media data with data from the news company's own platforms:

Oh, man, it would be great to have a virtual wall that I could just talk to and ask it to collect for me the viewing rates of a particular programme from particular years, or how many young viewers there were, or the time spent, or anything. (Manager 2, Data Set 1)

The managers also suggested that the organisations should develop structures that increase knowledge-sharing around data, enabling the entire staff to become more self-governed in their data practices. This vision thus touches upon the issue of how to best combine the data-informed management of individuals with the effective management of processes. From the perspective of journalists who only occasionally checked metrics, the usability imaginary had a rather down-to-earth mentality: they wished for technically stable systems and more 'journalistic' ways to name the metrics. Moreover, in the survey, less than half of the respondents reported that the analytics tools currently used within their organisations satisfied their needs.

The final vision that we identified deals with the *depth* of audience data, which seems to be in line with analytics companies' discourse on data as 'revealing' (Beer, 2019). In the newsrooms, depth refers to wishes expressed in relation to data quality and the related ability to draw more profound conclusions about users; hence, this vision deals with the domain of audiences. In their imaginations about the future, the interviewees saw increasing possibilities of 'enriching' data about content use with data from other sources, particularly from subscription systems. Furthermore, they also wished that legislation and platform companies would allow for more flexible collection and use of cross-platform data to obtain information on how their customers behave on social media platforms.

A related theme was gaining knowledge of the stories' significance for users or the impact that the news makes on audiences, starting with emotional reactions and reaching a wider societal impact, even negative effects. Manager 2 in Data Set 1 expressed that measuring 'the time spent' (Hendrickx et al., 2021; Milojevic, 2022) with the service was not enough. Instead, he wished to see a measurement for 'time *well* spent' that would indicate how valuable the news service has been to the reader. Many others also indicated that in a fruitful scenario, data would be used to better answer difficult 'why' questions:

Why did users decide to read a certain story and not another? Again, through negation, the survey responses indicated that the current use of analytics does not fulfil journalists' need to know more about the relationship that audiences have with journalism beyond single moments of use. In the survey, a respondent wrote:

Plain numbers are not enough. We need tools to figure out why a certain type of story has gained readers or why it has not. So, beyond the numbers, we need analysis. (Manager 71, Data Set 3)

Conclusion

In this article, we examined the components of data-informed news organisations as interpreted by journalism professionals. First, the pool of audience data given an informative role in Finnish newsrooms is vast and varied. Although we observed a pattern through our analysis (Figure 1), the basic elements appear somewhat shapeless in the professionals' interpretations. Hence, the undifferentiated category of *the general online audience* has the most discursive weight in decision-making, although the rise of the subscription-based business model positions the *paying audience* as the new ideal user. The latter indicates a move towards a more detailed and business-oriented audience understanding in the studied newsrooms.

Second, the above-described audience data inform the newsrooms in four main ways: (1) increasing the capabilities and performance of *individuals*, (2) streamlining work *processes*, (3) optimising *the news product*, and (4) engaging and controlling *the audience* (cf. Berndtsson et al., 2018; Cherubini and Nielsen, 2016: 10). Data-informed newsroom management in these domains happens through 'soft' means of anticipatory governance: steering, guiding, and nudging things forward in regard to next moves as well activities in the near future (see Flyverbom and Garsten, 2021; Petre, 2021). Despite some disagreement in the professionals' sensemaking, these domains were commonly recognised by managers, journalists, and supportive technologists, indicating an emerging order in the informative role of data.

Third, given the illumination provided by the analysis of the management domains regarding *where* and *what* data inform, the described visions, therefore, assess *how* audience data can maintain its role as a valid knowledge resource in newsrooms. For Finnish journalism professionals, audience data have the potential to (1) help newsrooms to better *predict* audience behaviour and business possibilities (2) if data use is *fair* for all parties and (3) data is offered to staff in a *usable* manner so that they can gain a (4) *deeper* understanding of their audiences and their work's impact. The visions seem to refer to both worries and hopes in uncertain times ahead (Flyverbom and Garsten, 2021: 8).

However, the professionals' visions cannot be labelled radical in that many of them were rather down-to-earth in nature. For example, imagining remained closely tied to existing technologies and certain success stories, such as editorial bots or Spotify. The visions were also clearly in line with the analytics companies' and consultants' imaginaries of data (Beer, 2019). The vision of fairness is the only key element that cannot be traced back to analytics companies' own imaginaries, rather to open data

movements (Baack, 2015), and the discussions on EU regulations and journalism ethics (Salonen, 2024). However, because all other visions bear resemblance to analytics companies' scenarios, and these companies operate strongly within the data-driven paradigm, negotiations around the future role of data in newsrooms are bound to continue, especially if journalism professionals continue with the view of placing data in a supportive and not leading role.

Continued negotiations, even tensions, are also a likely path because of the discrepancies we detected across the material. The visions pertaining to distant futures were more contested than present-focused sensemaking. That is, not all journalists were able to see how the gap between the present state and future possibilities for data-informed management could be closed. The journalists recognised the same patterns in the data-informed visions as managers and technologists did but took a more critical stance towards them (see Bunce, 2019).

Furthermore, there were internal contradictions between the different visions. For example, it might be difficult to negotiate between the ideas of gaining an even more detailed understanding of the audience from various data sources (the depth vision) while simultaneously respecting their privacy and the regulations (the fairness vision). In fact, we suggest that these tensions play a key role in how data-informed newsrooms are being constructed. It asks for a constant negotiation on organisational and individual levels of what the data 'say'. So, despite the extensive reliance on analytics programs as key *artefacts*, and the order created into data-informed newsroom management *practices* by the recognised domains, the *ideas* remained contradictory (cf. Hanitzsch, 2007). Hence, a coherent data culture (Ekström et al., 2022; Leino et al., 2023; Vulpius, 2023) is unlikely to be formed in the referred Finnish organisations.

We can also ask whether newsrooms should even aim for a shared data culture, as long as they have a shared journalism culture (Hanitzsch, 2007). Too solid framework could lead to a situation where the consensual and strict understanding of 'what the data say' would take over the individual and organisational possibilities to negotiate with data. The state of continued negotiation leaves opportunities for news professionals to evaluate whether data-based knowledge resonates with everyday experiences and journalistic values. Such evaluation implies a space for individual agency and anticipatory attunement with data (Koivunen et al., 2024).

Despite the illumination our findings provide, our study is limited by its focus on Finland only and the combination of three data sets originally intended to answer different (although connected) research questions. Nevertheless, we hope that the results regarding the domains of data-informed newsroom management and the related visions can provide insight for studies in other cultural contexts, especially in media environments where online subscriptions are surfacing as a business model.

Acknowledgements

The authors would like to thank Ritva Leino for fruitful discussion and comments at the early stages of the writing process, as well as the anonymous reviewers for constructive feedback.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the C.V. Åkerlund Media Foundation, Media Industry Research Foundation of Finland, Research Council of Finland (grant number 333503).

ORCID iDs

Laura Ahva  <https://orcid.org/0000-0001-6503-9772>

Margareta Salonen  <https://orcid.org/0000-0001-9635-2251>

Liisa Ovaska  <https://orcid.org/0000-0002-3305-5994>

Karoliina Talvitie-Lamberg  <https://orcid.org/0000-0003-0697-0385>

Notes

1. The project behind Data Set 1 aimed at answering the question: How are user analytics constructing what is regarded as news and journalism, and how is the audience seen as a result of this? The project related to Data Set 2 aimed at understanding how audience data is utilised and what its role is in newsrooms' decision-making processes among news professionals; and the project behind Data Set 3 studied how audience analytics are used and viewed in newspapers.
2. Approximately 600 journalists received the survey by email. The number is based on information from each newspaper's contact person.
3. Verbally in projects related to Data Sets 1 and 2, and in the project related to Data Set 3 by informing the participants through the survey template. All projects followed the [Finnish National Board on Research Integrity Guidelines \(2019\)](#). Accordingly, the projects were not pre-reviewed even though they dealt with human participants, as they did not pertain to minors or vulnerable groups.

References

- Ahva L and Ovaska L (2023) Audience metrics as disruptive innovation: analysing emotional work of Finnish journalism professionals. *Nordicom Review* 44(2): 152–171. DOI: [10.2478/nor-2023-0009](https://doi.org/10.2478/nor-2023-0009).
- Ananny M and Finn M (2020) Anticipatory news infrastructures. *New Media & Society* 22(9): 1600–1618. DOI: [10.1177/1461444820914873](https://doi.org/10.1177/1461444820914873).
- Anderson CW (2011) Deliberative, agonistic and algorithmic audiences: journalism's vision of its public in an age of audience transparency. *International Journal of Communication* 5: 529–547.
- Anderson C (2015) *Creating a Data-Driven Organization*. Sebastopol, CA: O'Reilly Media.
- Andrejevic M (2019) *Automated Media*. New York, NY: Routledge.

- Baack S (2015) Datafication and empowerment: how the open data movement re-articulates notions of democracy, participation and journalism. *Big Data & Society* 2(2): 1–11. DOI: [10.1177/2053951715594634](https://doi.org/10.1177/2053951715594634).
- Beer D (2019) *The Data Gaze: Capitalism, Power and Perception*. London: Sage.
- Berndtsson M, Forsberg D, Stein D, et al. (2018) Becoming a data-driven organisation. Presented at the 26th European Conference on Information Systems (ECIS2018), Portsmouth, UK, June 23-28, 2018.
- Bunce M (2019) Management and resistance in the digital newsroom. *Journalism* 20(7): 890–905.
- Cherubini F and Nielsen RK (2016) *Editorial Analytics: How News Media Are Developing and Using Audience Data and Metrics*. Oxford: Reuters Institute for the Study of Journalism.
- Christin A (2020) *Metrics at Work: Journalism and the Contested Meaning of Algorithms*. Princeton: Princeton University Press. DOI: [10.1515/9780691200002](https://doi.org/10.1515/9780691200002).
- Christin A and Petre C (2020) Making peace with metrics. *Sociologica* 14(2): 133–156. DOI: [10.6092/issn.1971-8853/11178](https://doi.org/10.6092/issn.1971-8853/11178).
- Chua S and Westlund O (2019) Audience-centric engagement, collaboration, culture and platform counterbalancing. *Media and Communication* 7(1): 153–165.
- Couldry N and Hepp A (2017) *The Mediated Construction of Reality*. Cambridge, UK: Polity.
- Ekström M, Ramsålv A and Westlund O (2022) Data-driven news work culture. *Journalism* 23(4): 755–772. DOI: [10.1177/14648849211052419](https://doi.org/10.1177/14648849211052419).
- Finnish National Board on Research Integrity Guidelines (2019) The ethical principles of research with human participants and ethical review in the human sciences in Finland.
- Flensburg S and Lomborg S (2023) Datafication research: mapping the field for a future agenda. *New Media & Society* 25(6): 1451–1469.
- Flyverbom M and Garsten C (2021) Anticipation and Organization: seeing, knowing and governing futures. *Organization Theory* 2(3). DOI: [10.1177/26317877211020325](https://doi.org/10.1177/26317877211020325).
- Flyverbom M, Leonardi P, Stohl C, et al. (2016) The management of visibilities in the digital age: introduction. *International Journal of Communication* 10: 98–109. <https://ijoc.org/index.php/ijoc/article/download/4841/1532>
- Hanitzsch T (2007) Deconstructing journalism culture. *Communication Theory* 17(4): 367–385.
- Hannila H, Silvola R, Härkönen J, et al. (2022) Data-driven begins with data. *Journal of Computer Information Systems* 62(1): 29–38. DOI: [10.1080/08874417.2019.1683782](https://doi.org/10.1080/08874417.2019.1683782).
- Hayes K and O’ Sullivan M (2024) Labouring the news: Management control strategies and work intensification in the digital newsroom. *Journalism* 25(6): 1346–1364. Available at: <https://doi.org/10.1177/14648849231174059>.
- Hellman H (2021) Lehtiketjut ja lehdistön ‘uusi normaali. *Keskittymisen mittaaminen ja sanomalehdistön keskittyminen Suomessa 2009–2019* 44(2): 92–116, Media & Viestintä.
- Hendrickx J, Montero E, Ranaivoson H, et al. (2021) Becoming the data-informed newsroom? *Digital Journalism* 9(4): 427–442. DOI: [10.1080/21670811.2021.1890170](https://doi.org/10.1080/21670811.2021.1890170).
- Hollifield CA (2019) Newsroom management. In: Vos TP, Hanusch F, Dimitrakopoulou D, et al. (eds) *International Encyclopedia of Journalism Studies*. Hoboken, NJ: John Wiley & Sons.
- Hupperz M, Gür I, Möller F, et al. (2021) What is a data-driven organization? Twenty-Seventh Americas Conference on Information Systems, 9-13th August, Montreal, Canada.
- Katzenbach C and Ulbricht L (2019) Algorithmic governance. *Internet Policy Review* 8(4): 1–18.

- Kitchin R (2014) Big Data, new epistemologies and paradigm shifts. *Big Data & Society* 1(1). DOI: [10.1177/2053951714528481](https://doi.org/10.1177/2053951714528481).
- Koivunen A, Nikunen K, Hokkanen J, et al. (2024) Anticipation as platform power: the temporal structuring of digital everyday life. *Television & New Media* 25(2): 115–132. DOI: [10.1177/15274764231178228](https://doi.org/10.1177/15274764231178228).
- Leino R, Koponen E, Kokko L, et al. (2023) *Asiakaskokemuksen Optimoinnin Opas*. Helsinki: Media-alan tutkimussäätiö. <https://optimointiopas.fi/>
- Lindén CG, Villi M, Lehtisaari K, et al. (2022) From clicks and time spent to value creation. In: Manninen VJE, Niemi MK and Rodge-Newman A (eds) *Futures of Journalism*. Cham: Springer International Publishing, 329–346. DOI: [10.1007/978-3-030-95073-6_21](https://doi.org/10.1007/978-3-030-95073-6_21).
- MacGregor P (2007) Tracking the online audience. *Journalism Studies* 8(2): 280–298. DOI: [10.1080/14616700601148879](https://doi.org/10.1080/14616700601148879).
- McAfee A and Brynjolfsson E (2012) Big data: the management revolution. *Harvard Business Review* 2012: 1–9.
- Milojevic A (2022) Datafication media and democracy: audience analytics and metrics in the Norwegian media organizations. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4245771.
- Napoli PM (2016) Special issue introduction: big data and media management. *The International Journal on Media Management* 18(1): 1–7. DOI: [10.1080/14241277.2016.1185888](https://doi.org/10.1080/14241277.2016.1185888).
- Napoli PM and Roepnack A (2018) Big data and media management. In: *Handbook of Media Management and Economics*. 2nd edition. London: Routledge, 410–421. DOI: [10.4324/9781315189918-26](https://doi.org/10.4324/9781315189918-26).
- Newman N, Fletcher R, Robertson GT, et al. (2022) Reuters institute digital news report 2022. In: *Reuters Institute for the Study of Journalism*. Oxford: University of Oxford. <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2022>.
- Petre C (2021) *All the News That's Fit to Click: How Metrics Are Transforming the Work of Journalists*. Princeton: Princeton University Press.
- Porlezza C (2023) The datafication of digital journalism: a history of everlasting challenges between ethical issues and regulation. *Journalism* 25(5). DOI: [10.1177/14648849231190232](https://doi.org/10.1177/14648849231190232).
- Rolandsson T, Widholm A and Rahm-Skågeby J (2022) Managing public service: the harmonization of datafication and managerialism in the development of a news-sorting algorithm. *Digital Journalism* 10(10). DOI: [10.1080/21670811.2022.2119151](https://doi.org/10.1080/21670811.2022.2119151).
- Salonen M (2024) *Revising the understanding of gatekeeping theory: Factors and practices of post-publication gatekeeping*. JYU Dissertations. Jyväskylä: University of Jyväskylä 776. Available at: <http://urn.fi/URN:ISBN:978-952-86-0142-5>.
- Salonen M, Ehrlén V, Koivula M, et al. (2023) Post-publication gatekeeping factors and practices: data, platforms, and regulations in news work. *Media and Communication* 11(2): 367–378. DOI: [10.17645/mac.v11i2.6486](https://doi.org/10.17645/mac.v11i2.6486).
- Sirén-Heikel S, Kjellman M and Lindén (2022) *C-GJournal of the Association for Information Science and Technology* (2022) *At the crossroads of logics: Automating news work with artificial intelligence* 74(3): 354–366. DOI: [10.1002/asi.24656](https://doi.org/10.1002/asi.24656).
- Svensson J (2023) Logics, tensions and negotiations in the everyday life of a news-ranking algorithm. *Journalism* 24(7): 1518–1535.

- Tandoc EC (2019) *Analyzing Analytics: Disrupting Journalism One Click at a Time*. London: Routledge. DOI: [10.4324/9781138496538](https://doi.org/10.4324/9781138496538).
- Tenor C (2023) Metrics as the new normal – exploring the evolution of audience metrics as a decision-making tool in Swedish newsrooms 1995-2022. *Journalism* 25(5). DOI: [10.1177/14648849231169185](https://doi.org/10.1177/14648849231169185).
- Vulpus J (2022) The role of audience data and metrics in becoming ‘audience-centred’. In: Manninen VJE, Niemi MK and Rodge-Newman A (eds) *Futures of Journalism*. Cham: Springer International Publishing, 347–363.
- Vulpus J (2023) “We need to think about their real needs”: examining the auxiliary work of audience-oriented intralopers in news organizations. *Digital Journalism*. DOI: [10.1080/21670811.2023.2288245](https://doi.org/10.1080/21670811.2023.2288245).
- Welbers K, van Atteveldt W, Kleinnijenhuis J, et al. (2016) News selection criteria in the digital age. *Journalism* 17(8): 1037–1053. DOI: [10.1177/1464884915595474](https://doi.org/10.1177/1464884915595474).

Author biographies

Laura Ahva (PhD) is an Associate Professor in journalism studies at Tampere University in Finland. Her research interests include journalistic work, audience studies, participatory media, and practice theory. She has published in journals such as *Digital Journalism*, *Journalism Studies*, and *Journalism Practice*.

Margareta Salonen (PhD) is a Postdoctoral Researcher at the University of Jyväskylä, Finland. Her work focuses on the triadic power relations between news media, audiences, and social media platforms through the lens of (post-publication) gatekeeping. Her work also covers datafication of news work, journalism ethics on social media platforms and the role of visibility and regulations in social media journalism. She has published a co-authored book with Palgrave Macmillan and articles in journals such as *Journalism Practice*. She is a member of the Finnish Press Council (2021–2026) and currently the PI of *Ephemerality and Visuality in Social Media News Content* (2024–2027).

Liisa Ovaska is a Doctoral Researcher at Tampere University in Finland. Her dissertation project focuses on the journalists’ and the news users’ reflections on the usage of audience data in journalism. She has published her work, for example, in *Nordicom Review*. Ovaska holds a master’s degree (MA) in social sciences (in journalism) from Tampere University.

Karoliina Talvitie-Lamberg (PhD) is an Assistant Professor in communication and media research at the University of Jyväskylä in Finland. Her recent research centers on digital disconnection, belonging in datafied every day, and related experiences of vulnerability. She has published in journals such as *New Media and Society*, *Computers in Human Behavior*, *Media and Communication*, *Television & New Media*, *Media Theory*, *Surveillance*, and *Society*.