



COMMUNICATIVE FIGURATIONS | WORKING PAPER | No. 18

ISSN 2367-2277

Wiebke Loosen

**Four forms of datafied journalism
Journalism's response to the datafication of society**



Forschungsverbund „Kommunikative Figurationen“ | Research Network “Communicative Figurations”
Universität Bremen | University of Bremen
ZeMKI, Zentrum für Medien-, Kommunikations- und Informationsforschung
Linzer Str. 4, 28359 Bremen, Germany, E-mail: zemki@uni-bremen.de
www.kommunikative-figurationen.de | www.communicative-figurations.org

Wiebke Loosen (w.loosen@hans-bredow-institut.de)

Wiebke Loosen is a Senior Researcher for journalism research at the Hans-Bredow-Institut for Media Research in Hamburg as well as a Lecturer at the University of Hamburg. Her major areas of expertise are the transformation of journalism within a changing media environment, theories of journalism, methodology and constructivist epistemology. Wiebke Loosen's current research includes work on the changing journalism-audience relationship, datafied journalism, the emerging 'start-up culture' in journalism, as well as algorithms' 'journalism-like' constructions of public spheres and reality.

Working Paper No. 18, March 2018

Published by the „Communicative Figurations“ research network, ZeMKI, Centre for Media, Communication and Information Research, Linzer Str. 4, 28359 Bremen, Germany. The ZeMKI is a research centre of the University of Bremen.

Copyright in editorial matters, University of Bremen © 2018

ISSN: 2367-2277

Copyright, Electronic Working Paper (EWP) 18 - Four forms of datafied journalism. Journalism's response to the datafication of society. Wiebke Loosen, 2018

The author has asserted his moral rights.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means without the prior permission in writing of the publisher nor be issued to the public or circulated in any form of binding or cover other than that in which it is published. In the interests of providing a free flow of debate, views expressed in this EWP are not necessarily those of the editors or the ZeMKI/University of Bremen.

Four forms of datafied journalism

Journalism's response to the datafication of society

1 Introduction

Grand narratives used to describe the state of society tend to sensitize us to transformation processes that have the potential to lead to a vastly different perception of the social world. Many of these narratives we hear today tell stories of data or even *big data*. Here we find notions of big data as “a revolution that will transform how we live, work and think” (Mayer-Schönberger and Cukier, 2013), considerations on “the society of data” (Süssenguth, 2015), and the “datafied society” (Schäfer and Van Es, 2017). This illustrates that the idea of datafication is not only used to describe how digitization is transforming our media environment: in a much more fundamental way it is actually a story about how numerical data have come to represent, and at the same time influence, social reality (Van Dijck, 2014).

This paper situates the datafication of journalism in relation to society's datafication. This is understood as a useful step in our theorizing of three interrelated elements: journalism, data, and social reality.¹ By bringing these three elements together, I consider journalism as an ideal example to understand how datafication shapes and transforms a social domain and how it influences public communication. This allows us to not only better comprehend journalism's present transformation towards a more data-based, algorithmed, metrics-driven, or even automated practice, but, to consider this transformation as a reflexive process: a process that is at the same time part of a changing media environment and is journalism's response to – as well as an act of encouraging – the datafication of society. It is important, therefore, to recognize that journalism does not simply work *with* media technologies but is operating *within* a constantly changing media environment. We might frame this reflexive process as the transition towards a *datafied journalism within a datafied media environment of a datafied society*.

It can be argued that datafication represents one of the most pressing and recent trends of a changing media and communicative environment (Hepp and “Communicative Figurations” research network, 2017). What makes journalism an ideal example for deepening our understanding of datafication is that journalism has always been a core domain of public communication - adapting itself to the changing media and communicative environment it operates within (Loosen, 2015). The phenomenon of datafication itself demands that this example automatically extends beyond journalism as it comes with transformations that, among other things, facilitate journalism-like performances and services outside of what used to be described as the field of journalism. Thus, on the one hand, processes of datafication appear in journalism as if seen through a magnifying glass; on

¹ I owe this appreciative description of my approach to an unknown reviewer who reviewed a previous version of this text.

the other hand, however, they go far beyond: they fundamentally transform society's communicative foundations and come with far reaching consequences for the construction of public spheres.

In what follows, I begin by outlining the development of prevalent thought relating to the datafication of society, the fundamental ideas behind it and the evidence thus far presented to demonstrate these claims about the datafication of society. Afterwards, I present a heuristic consisting of four forms of datafied journalism: data-based journalism, algorithmed journalism, automated journalism, and metrics-driven journalism. The substrate of this discussion is summarized by positioning the figure of 'the algorithm' at the center of my concluding remarks.

2 The datafication of society

Journalism's role is a societal one: it is legitimized and (in Western democracies) legislated for as if it provides a particular service to society. This is the reason why journalism research has comparatively many references to social and societal theories (Löffelholz and Rothenberger, 2016; Steensen and Ahva, 2015). A lot of empirical journalism research, however, does not explicitly address the subject at the level of society - and in most cases, this is neither necessary nor useful. However, if we want to develop a deeper understanding of journalism's transformation towards a *datafied journalism - within a datafied media environment - within a datafied society* it will be necessary to situate this transformation within a broader societal context, we need to consider the datafication of journalism in relation to the datafication of society.

The underlying direction of thought interested in the entanglement of and the reflexive relationship between media change and societal change is reflected in a variety of social theories, theories of society, and grand theoretical approaches or macro concepts such as 'network society' (Castells 2000), 'communication society' (Münch 2002), or in systems theory (in the sense of Luhmann, 2000, 2012), and (deep) mediatization theory (Couldry and Hepp, 2017). As different as the two latter approaches in particular are in detail, they both exercise restraint towards linear causal relationships in favor for non-linear reflexive ones, they are constructivistically founded (Hepp et al., 2017), and they share a simple but at the same time momentous idea that this paper is also attempting to foreground: They all assume that social domains/fields/systems, that the social world/social reality/society is, in essence, based on (or: operate on the basis of) communications, on *communicative constructions*. These communications are facilitated by certain (electronic dissemination) media of communication that have been introduced to society at certain points in history (language, writing, printing press, computers - as systemized by Luhmann, 2012; Baecker, 2007a, b). As a consequence, specific communicative conditions are shaped by the media that are available at specific points in a media ecology's evolution, societal change and media change are intrinsically interwoven.

From our present perspective, the current changes in the media environment are most starkly demonstrated by computerized communications and the ascent of the internet – which itself comprises a wide range of media offerings and various modes of communication. Niklas Luhmann's oft-cited reflection on "the reality of the mass media" (2000), that "whatever we know about our society, or indeed about the world in which we live,

LOOSEN: FOUR FORMS OF DATAFIED JOURNALISM

we know through the mass media” still sensitizes us to the fact that our knowledge of the world and how we make sense of it is highly dependent on mediated communication, despite, by virtue of its focus on mass media, maintaining less credence in a landscape characterized by digital media technologies.

This leads us to an important point: datafication is only one of many trends observable within our changing media environment. In all, five trends can be identified to characterize the contemporaneous transformations of our media environment (Hepp and “Communicative Figurations” research network, 2017): besides *datafication* these are the *differentiation* of media devices; their increasing *connectivity* through the infrastructure of the internet; the *omnipresence* of digital media as demonstrated by mobile communications; the *rapid pace of innovation* in terms of content, platforms, and (mobile) devices. None of these trends should be seen as discrete phenomena; rather, they are intimately linked, and together they characterize with clarity the changes our media environment is currently going through. As broad as they may be, they only really offer an initial understanding of media-related change that requires further investigation with regard to individual phenomena and specific social domains.

Datafication, however, can be regarded as the most recent “wave” in an era of rapid media change and one crucial characteristic of the current mediatization of our social world (Couldry and Hepp, 2017: 52). With media’s digitization they become less mere means of communication but, means of generating data as well (Breiter and Hepp, 2018). The digitalization of our (media) environment – and the digital traces and big data that accrue with living in such an environment (Deuze, 2012) – turns many aspects of reality, the world, the social, social life, and social action into computerized data – data that is, to various ends, aggregated and processed algorithmically (Kitchin, 2014; Mayer-Schönberger and Cukier, 2013; Van Dijck, 2014).

In fact, algorithms do not exist in a vacuum, they do not amount to anything without the data they process and it is exactly this “dyad of big data and algorithms [that] can enable new cultural and social forms” (Uricchio, 2017: 126). Algorithms and data play a role in fundamental spheres such as “access to information (Google) and the social world (Facebook), [...] in finance (algorithmic trading) and governance (from predictive policing to NSA-style parsing of vast troves of data)” (Uricchio, 2017: 128). It seems to be the case that the notion of datafication, itself a process, is again based on diverse data-related processes: data generation (digital traces, big data), the processing of these data by algorithmic means (algorithms), and automation enabled by the two former processes (e. g. in the case of journalism: automated content production).

In essence, these discourses hark back to the fundamental question about what changes in a society when communication also includes besides humans machines and algorithms (Esposito, 2017). Backgrounding this question is the normative query on the extent to which we wish to view our society as *computable* and also the extent to which we wish to have it *computed*, not to mention which observations and decisions we want computers and algorithms to make. This touches on further questions on the epistemological status of big data (boyd and Crawford, 2012; Kitchin, 2014), algorithmic accountability (Diakopoulos, 2015), and an ethics of algorithms (Ananny, 2016), illustrating the idea that computation is not only about binary code, but about constructing social reality:

“algorithms do not simply accelerate commerce, journalism, finance, or other domains – they are a discourse and culture of knowledge that is simultaneously social and technological, structuring how information is produced, surfaced, made sense of, seen as legitimate, and ascribed public significance” (Ananny, 2016: 98).

Datafication, however, does not follow a linear trajectory nor does it take place at the same speed or in similar ways to developments in different social domains: datafication leads to a variety of consequences and manifests itself in different ways, for instance, in politics (Madsen et al., 2016) than it does in the financial world (Knorr Cetina and Reichmann, 2015), and again, in the realm of education (Williamson, 2017). However, what all social domains have in common is that we can assume that they will increasingly rely on an ever more diverse and greater amount of data in their (self-) sense making processes. We can see the same developments occurring in the scientific field through its exploration of a set of new digital methods with which to study society, methods that use digital traces that users leave behind in their use of contemporary media; behaviors that are intrinsic to everyday life (Rogers, 2015; Schäfer and Van Es, 2017). So what seems to be taking place is a “gradual normalization of *datafication* as a new paradigm in science and society” (Van Dijck, 2014: 198, emphasis in original).

Journalism, in the broadest sense of the term, is meant to observe society, and presently, it needs to be able to make sense of these developments to fulfil its societal functions and to adapt itself to a datafied media environment (Loosen et al., 2017). This transition towards a datafied society is reflected in journalism in the following ways: Datafication has become an *object of reporting* that reflects the societal relevance of the phenomenon and has experienced a boom in recent years and is considered an appropriate avenue of inquiry for almost all editorial departments including politics, economics, sports and so on. Datafication refers to an *entanglement of work practices* that stimulates a computational turn in journalism’s essential character. And, as a broad societal trend, datafication facilitates functional *equivalents of journalism* by non-journalistic services and providers such as search engines, social networks and content producing technology companies.

These processes illustrate how datafication touches journalism on multiple levels as well as reaching beyond the field: it represents a new topic in journalism’s agenda influencing its public perception and awareness; it affects the means and practices by which journalism operates and observes society, and, by changing the environment journalism is operating within, it touches the functional level of journalism and its relations with this environment – including the one to its audiences.

3 Four forms of datafied journalism

Journalism is a genuine media phenomenon and as such is inextricably sutured to the (transformation of the) wider media environment. This reflexive relationship affects how journalism observes the world, how it reflects on itself, how it is produced and (re-) presented in different forms, how it is distributed, and how it is ultimately used by audiences (for a historical perspective see Birkner, 2012, for latest perspectives Boczkowski and Anderson, 2017). If this media environment becomes a datafied media environment, this will also have an impact on journalism.

LOOSEN: FOUR FORMS OF DATAFIED JOURNALISM

To refine this general observation, I suggest a heuristic of *four forms of datafied journalism* that is built around the entanglement of *the process of news production and consumption* (represented by the inner circle in Figure 1) and *the process of datafication* (represented by the outer circle in Figure 1). The intention to relate these circles is to contextually link the changes in the communicative conditions of an increasingly datafied society in general (datafication) and of journalism in particular (datafied journalism).² All processes, news production and consumption as well as datafication, can be further differentiated:

First, the ‘circuit of news’³ as the circular process of news production and consumption consisting of common components of communication processes that characterize the news production process (Domingo et al., 2008): *observation*⁴, *production*, *distribution*, and the *consumption* (interpretation/understanding/sense making) of (journalism’s) communicative offerings by an audience. However, these components represent complex processes in themselves and are by no means linear; in their essence they constitute the basic circular relationship between journalism and its audience (see section metric-driven journalism): Every journalistic observation of the world (in order to turn these observations into ‘the news’) is already informed by journalists’ expectations of what audiences may want and audiences consumption habits are, in turn, always already informed by expectations about what journalism will or should deliver (Loosen and Schmidt 2012).

Second, datafication is considered as a context for the whole circuit of news: news production and consumption occurs in an increasingly datafied and data-driven media environment. Again, datafication itself can be understood as a process that includes other data-related processes building upon each other; these are: *data generation* (leading to an increase in data availability), *algorithmic data processing* (making the algorithm and its operational modes an object of investigation), and automated production (facilitated by the availability of data, processing power, and automation of processes by computational means) (see Figure 1). While the distinction between these three processes is a conceptual one, they are in fact, mutually dependent and by this have a close interrelation: algorithmic data processing and automation can be seen as the follow up processes to data generation; algorithms process data and in turn automate the entire procedure: the automation of content relies on the first two processes. Or, put differently: data’s increasing availability demands increasingly powerful algorithmic processing – together, both processes enable automation. These three processes of datafication are, of course, not only relevant for journalism, but stimulate comparable occurrences in other domains.

² The idea is not to explain the differences and similarities between forms of datafied journalism better than others already have done (Coddington, 2015) or to ultimately distinguish them from each other, but, while drawing on such attempts, to situate developments towards an increasingly datafied form of journalism in relation to datafication as an overarching process of media and societal transformation.

³ In loose analogy to the ‘circuit of culture’ (du Gay, 1997; du Gay et al., 1997).

⁴ *Observation* is understood here as the mode/way (e.g. through the lens of news values as news construction principles) in which journalism observes the world. Since observation inherently involves selection, it is in itself a process of reducing complexity, I do not, like Domingo et al. (2008), distinguish selection from observation.

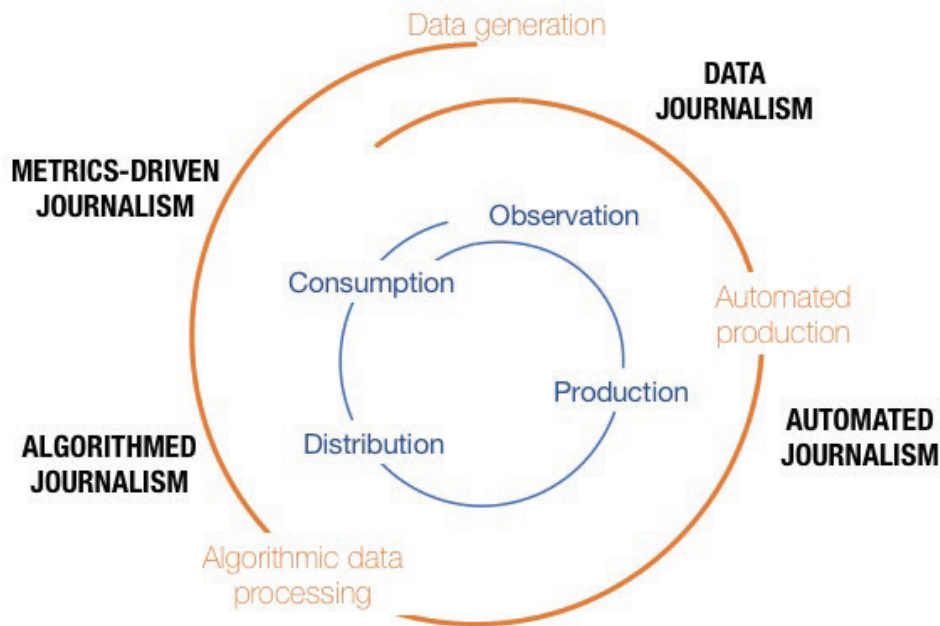


Figure 1. The circuit of news, processes of datafication and four forms of datafied journalism.

In fact, we can consider that most of the research related to journalism’s datafication is organized around the stages of news production and to one or the other data-related process (Lewis and Westlund, 2015).

Reflecting the circuit of news in the light of datafication, we can highlight the most striking and challenging transformations journalism faces today and identify four forms of datafied journalism (see Figure 1):

- **Data journalism:** representing the emergence of a new style of reporting based on *sourcing* ever more (publicly) available data sets (Loosen et al., 2017);
- **algorithmed journalism**⁵: the increasing influence algorithms have on the *distribution* of journalistic content as they – having been ‘built in’ to various services and platforms – make journalism-like judgements such as selecting, prioritizing, and defining what is relevant (Gillespie, 2014);
- **automated journalism:** the increasing amount of content that is being *produced automatically* and by means of technologies being developed by providers of automated content solutions that do not consider themselves journalistic organizations (Dörr, 2016; Graefe, 2016);
- **metrics-driven journalism:** the varied attempts to *make sense of an ever-growing amount of audiences’ digital traces* with the potential to influence decision making processes at all stages of the news production process (Tandoc and Thomas, 2015).

It is obvious that these four forms are not just examples for a datafied journalism, but

⁵ In the context of the four forms of datafied journalism, which I distinguish here, I prefer to use the term “algorithmed journalism” instead of “algorithmic journalism” to emphasize that the influence of algorithms can refer to the distribution of all sorts of journalistic materials that are does not necessarily have to be influenced by algorithmic data processing. Algorithmic journalism, on the other hand, is generally used either to characterize pars pro toto a computational turn in journalism (Anderson 2012), or specifically for automated journalism (Dörr 2016).

LOOSEN: FOUR FORMS OF DATAFIED JOURNALISM

rather, they address different, yet interrelated, levels of journalism in a datafied society affecting (1) *journalism's way of observing the world* and constructing the news by the means of data (data journalism); (2) the *distribution and circulation of journalism's output* within an environment that is shaped by algorithms and their underlying logic to process data (algorithmed journalism); (3) the very core of journalism's performance in *facilitating the automation of content production* (automated journalism); (4) what is *understood as newsworthy* to ever more granularly divided audience segments (metrics-driven journalism).

To bring it all together, we can see that datafication as a larger societal transformation process touches journalism not only at the level of the basic stages of news production and consumption, it actually goes beyond a quantitative or computational turn that affects journalism at its core (Anderson, 2013; Coddington, 2015) – and beyond what a fruitful “sociotechnical emphasis in journalism studies” (Lewis and Westlund, 2015: 20) is able to reveal. Through the follow-up processes of algorithmic data processing as a means to radically *reduce the complexity* of ever-increasing amounts of data via automation and, at the same time, as a means to *increase complexity* by producing more data – it also, first and foremost, re-shapes the communicative environment journalism operates within. Journalism is, then, profoundly affected and increasingly driven by datafication, and, as an early adopter of digital media technologies and (big) data as a source, is at the same time a driver of this process.

Data journalism

Today's data journalism is mainly contextually situated within many different forms of data-driven journalism practice and is in many ways related to what is referred to as the ‘quantitative turn in journalism’. Coddington (2015), for instance, differentiates between *computer-assisted reporting*, serving as the historical foundation the more recent forms data journalism and computational journalism are based on. However, for the sake of our argument in this paper, it is not so much decisive in terms of the extent to which data journalism is similar to and different from other data-driven journalism practices. What is crucial here is that it usually builds on (large) sets of (digital) data as ‘raw material’ that is subjected to some form of (statistical) analysis in order to identify and tell stories (Coddington, 2015; Royal and Blasingame, 2015). Contemporary data journalism relies on the (increased) availability of data (sets) to expand the repertoire of sources for journalistic research and for identifying stories. Data journalism represents the emergence of a new journalistic sub-field that is co-evolving with the datafication of society – a logical step in journalism's adaptation to the increasing availability of data traveling in tandem with datafication as a data generating process.

One recurring finding in content-related research on data journalism is that it exhibits a ‘dependency on pre-processed public data’ (Tabary, Provost and Trottier, 2016: 75; cf. also Borges-Rey, 2017; Young, Hermida and Fulda, 2017) from statistical offices and other governmental institutions, for example. This is also true for data-driven pieces at an award-worthy level: In our own analysis of projects nominated for the Data Journalism

Awards⁶ through the years 2013 to 2016, we observed a dependence on data from official institutions or other non-commercial organizations such as research institutes, NGOs and so on; data that are publicly available or is available, at least, on request (Loosen et al., 2017a). This illustrates, on the one hand, that journalism is making sense of the increased availability of data sources, but on the other, that it also relies heavily on these data: the share of self-collected, scraped, leaked, and requested data is substantially smaller. Nonetheless, data journalism has been continually linked to investigative reporting (Parasie, 2015; Royal and Blasingame, 2015), which has ‘led to something of a perception that data journalism is all about massive data sets, acquired through acts of journalistic bravery and derring-do’ (Knight, 2015). Recent cases such as the ‘Panama Papers’ have contributed to that perception (<https://panamapapers.icij.org>). This example demonstrates that some topics of global concern need to be based on international data that require transnational collaborations between various media organizations.

Data journalism’s reliance on certain sources has an effect on the topics it may or may not cover. As a result, data journalism can neglect those social domains for which data are not regularly produced or accessible. Content analyses identified a preponderance of political pieces (Tandoc and Oh, 2017) and considerable coverage of societal (Knight, 2015; Young, Hermida and Fulda, 2017), business (Parasie and Dagiral, 2013), and health issues (Young, Hermida and Fulda, 2017). Yet, most of this empirical research is spatially and temporally restricted and typically focuses on particular case studies. An international comparative study that relates data availability and accessibility in different countries to topics covered by data reporting is still absent from the literature but could, once carried out, shed light on which social domains and topics are covered by which analytical methods and on the basis of which data sources. Such an approach would also give insight to the opposite: the blind spots in data-driven coverage due to a lack of (available) data sources.

However, we can expect data journalism’s relevance and proliferation to co-evolve alongside the increasing datafication of wider society. The more the social domains that journalism is supposed to observe are themselves datafied – that is, the more their (self-) sense making and social construction rely on data, the more journalism itself needs to be able to make sense of data to fulfil its function. Additionally, data-driven journalism’s growing relevance may set incentives for other social domains to produce or make more data available (to journalists). In the event of this happening, we are likely to see the co-evolution of a ‘data PR’, that is, *data-driven public relations* produced and released to influence public communications for its own purposes.

Algorithmed journalism

The increasing availability of data brings with it the need to process that data. This is where the complexity reducing capacity of algorithms comes into play. The term ‘algorithmed journalism’ in our heuristic refers to the increasing influence algorithms have in the distribution of journalistic content. This is not to say that algorithms do not play a role in datafied journalism in general, in fact, the interference of algorithms goes far beyond

⁶ The Data Journalism Award is a prize awarded annually by the Global Editors Network; <https://www.globaleditorsnetwork.org/about-us/>.

LOOSEN: FOUR FORMS OF DATAFIED JOURNALISM

pure distribution (Napoli 2014). Having been built into various services and platforms, algorithms are even able to make journalism-like judgements such as selecting, prioritizing, and defining what is relevant. Gillespie (2014: 168) refers to these types of algorithms – those capable of reproducing the functions of journalism – as ‘public relevance algorithms’.

The influence social media platforms like Facebook and Twitter have on the automation of journalistic judgements is characterized by the phenomenon of the so-called “platform press” that has somehow “reengineered journalism”. Bell and Owen (2017) argue that “publishing is no longer the core activity of certain journalism organizations” as those responsibilities move to social media platforms that rely on algorithms. Behind observations of this kind stands the conviction that algorithms are “now a key logic governing the flows of information on which we depend” (Gillespie, 2014: 167). In any case, algorithmed journalism as one of the four forms of datafied journalism suggested here stands for the idea that the distribution and targeting of journalistic content have become closely entangled with algorithmic data processing on social media platforms to the extent that we could alternatively refer to the practice as *platform-driven journalism*.

In the realm of social media platforms, the communicative conditions differ from those in the mass media. New kinds of algorithmic personalization apply, for instance, when algorithms reflexively calibrate their selection and sorting capacity according to an individual user’s prior behavior (or, more correctly: a user’s digital traces) and, in their last analysis, generate a ‘deep personalization’. This type of personalization is considered by some to be dysfunctional for the production of a genuine public sphere, it represents a ‘new media logic’ (Couldry and Turow, 2014: 1711). In this regard, the Facebook news feed, for example, does not actually address a dispersed mass audience but addresses individual users by taking advantage of various parameters to calculate configurations of users’ digital traces and those of their “friends” resulting in individual notifications being displayed more prominently – or not at all (Bakshy, Messing and Adamic, 2015; DeVito, 2017). As with journalism, this is a process of selecting and producing of what is more or less relevant. The underlying decisions, however, proceed from divergent presumptions (general-social vs. specific-individual relevance) made on the basis of data collected from individual users and the audience as a whole (aggregated knowledge about ‘the audience’ vs. individual user data) combined with various assumptions, or constructions, of audiences (citizens requiring information vs. users requiring satisfaction, for example). With this in mind, algorithmically constructed offerings and mass media offerings follow different modes of communication or frames of relevance that we can locate at the poles of generalization (the public) and personalization (the user).

However, as regards to the *social* significance of news, both frames of relevance are not necessarily mutually exclusive. The social significance of news is exactly what plays a primary role in social media’s operations, when it is namely a question of whether or not the items recommended within your own network, are shared or ‘liked’. This indicates that we should not consider professional journalism and algorithms as two separate agents, but as interwoven with and relating to each other. Journalism reports on selected events or it produces news items for a generalized audience. Wider distribution to what is, as a rule, an individualized audience occurs when algorithms process already existing information (not necessarily, although quite often, journalistically generated). We can describe this as algorithmically operationalizing news factors from the viewpoint of users. (Loosen and

Scholl, 2017)

Brought together, this illustrates that the “communicative role of algorithms is clearly a massive social phenomenon with many complex consequences” (Esposito, 2017: 250) - and there is little doubt that this phenomenon plays a crucial role in communication, public communication in particular, for which journalism used to maintain a gatekeeper monopoly. Meanwhile, algorithms are, due to their increasing entanglement with various kinds of communicative processes and practices, considered to have a similar capacity for the production of public spheres and constructions of reality as traditional journalism has had throughout its history. This is the principal characteristic of ‘algorithmed journalism’ and it goes well beyond any idea of ‘platform-driven journalism’. It is exactly these different orientations or frames of relevance for journalism and social media as well as their underlying mechanisms of self-reinforcement that form the current fears of self-reinforcement mechanics that lead to so-called filter bubbles – no matter how justified or unjustified they may be (Zuiderveen Borgesius et al., 2016).

Automated journalism

We can think of automated journalism as being at the most sophisticated level of a datafied journalism. Data does not represent, as it does in data journalism, merely the sources used to identify and tell stories, but is in fact the decisive element in the “algorithmic processes that convert data into narrative news texts with limited to no human intervention beyond the initial programming” (Carlson, 2015: 417). Even if it comes with the promise of respite from routine work (Graefe, 2016; Lindén, 2017), automated content production is one logical consequence of datafication and at the same time a driver of it. It simultaneously reduces and produces complexity as the technology, on the one hand, ‘makes sense of ever – more available data by computational means, while, on the other hand, generates more and more content on the basis of increasingly sophisticated computation.

This technology, facilitated by advances in the field of automated content production driven by fields such as artificial intelligence and natural language generation, affects journalism at its very core: the production of news. Montal and Reich (2017: 829) argue that automated journalism (also referred to as “robot journalism”; Clerwal, 2014; Van Dalen, 2012) comes with “significant practical, sociopolitical, psychological, legal and occupational implications for news organizations, journalists and their audiences”.

Developments in the field make us aware of the fact that journalistic content is no longer produced by media organizations only. One of the main players in the field of automated content production is Automated Insights, a technology company whose customers are not only media organizations but organizations in other industries such as e-commerce or financial services (<https://automatedinsights.com/use-cases>). The same technology used to produce journalism is also being used to produce product descriptions or market reports, for example. This doesn’t need to be a problem in itself: we also use the same language when producing journalism and public relations. However, this raises crucial questions about “who should be acknowledged as the author of automated items” (Montal and Reich, 2017: 830) and on the underlying methodologies used and data sources exploited in the content generation process. The “attribution policy for algorithmic content generation” Montal and Reich (2017: 843-844) devised to ensure algorithmic transparency, suggests that those obligations (and rights) usually applied to journalism and media organiza-

LOOSEN: FOUR FORMS OF DATAFIED JOURNALISM

tions will eventually expand their reach to include this new form of content and, moreover, to also take into account non-media organizations such as technology firms and non-journalists such as programmers and data scientists, practitioners that are increasingly participating in journalism-like or ‘pure’ journalistic services. It is, however, doubtful to what extent and by what means a “full disclosure” of the “algorithmic methodology” (Montal and Reich, 2017: 843) can be achieved – or if in the end it will have any relationship to or similarity with journalism in general. It may become, rather, a question of institutional trust and legitimization that, by including a functional proportion of critique and observation, societally secures this new form of information/knowledge production as part of the societal information ecosystem. However, these developments and regulatory attempts to achieve algorithmic accountability (Diakopoulos, 2015) are hard indicators of the fact that we take algorithms increasingly seriously as relevant actors shaping society’s communicative foundations.

Automated journalism has an inherent paradoxical nature: on the one hand, it reduces complexity by automatically making sense of the ever-growing availability of data, while on the other, it increases it by producing an increasing volume of content – content that otherwise may not have been produced at all. In light of this contradiction, automated journalism appears to resemble an *algorithmically reloaded information overload* illustrating the observable relationship between a simultaneous reduction of and increase in complexity. Natural language processing solves the problem of making sense of data, while simultaneously generating a consequential ‘problem’ that we can ultimately refer to as ‘understanding’. This problem is again, given the amount of data that is being produced and processed, solved by computational means. This means, in a final analysis, that natural language processing operates in a cycle of automated content production and automated content understanding.

Metrics-driven journalism

Audience metrics in journalism is probably the area for which an increased drive towards datafication is most apparent. The cycle of news production and consumption, of supply and demand, can be understood initially as the basic circular relationship between journalism and its audience; it is, in essence, about the communicative offerings journalism provides that are in turn accepted by an audience (Görke and Scholl, 2006; Loosen and Schmidt, 2012). However, there is a more pertinent line of empirical inquiry related to how journalism’s production process relates to (expectations related to) consumption and whether or not this process changes in the light of a more metrics-driven practice (Wang, 2017).

Today this question is discussed with an eye towards the current media landscape and the increasingly sophisticated opportunities to gather and analyze the digital traces of audiences’ activities. Consequently, variations on the question of “how [...] [the] transformations in newswork [that are] intersecting with changes in the monitoring of reader behavior and new technologies of audience measurement” (Anderson, 2011: 550) are being interrogated with increasing vigor and are evolving in parallel with newly available digital traces left by users on news websites, social media, or comment sections and fueling the rise of a whole new ‘audience measurement industry’ (Napoli, 2010; Nelson and Webster, 2016). This “calculated measurements” (Passoth, Sutter and Wehner, 2014) are reshaping

the relationship of providers and audiences in the whole field of cultural production.

Digital traces contribute to a new level of audience visibility; audiences' utterances represent a unique source of qualitative information. Through direct contact or by observing follow up communication to their stories, journalists learn about their audiences' opinions, questions, doubts and critique, and about their attitudes towards and expectations of journalists and their work (Hanusch and Tandoc, 2017) – in turn, audience members become more visible to other users, too. Primarily, however, the digital media environment contributes to the datafication of audience behavior through quantitative data (Ang, 1991) and subsequently contributes to a “data-driven audience understanding” (Wang, 2017: 2). For instance, audience members' digital traces – like click rates and social media analytics – reveal information on news preferences, evaluation, and engagement (Cherubini and Nielsen, 2016). These data are also ‘fed’ to algorithms that are implemented on different platforms, apps, and services, selecting, filtering and ranking news items according to calculated individual user preferences (Tandoc, 2014). Through these implementations, we can also see the tight bond between algorithmed and metrics-driven journalism and the particular relevance social media platforms and their algorithms have to the distribution, visibility, and consumption of news. Algorithms are binding the production – distribution – consumption cycle tighter than ever and assist in evaluating various forms of user data to produce media products that will have the widest possible reach – leading to a kind of algorithm-driven recursive loop. It is an open question to what extent journalism's fundamental orientation of giving the public what it *should* know remains for professionals.

Consequently, there is no doubt that the still new and increasingly sophisticated forms of audience measurement have an influence on news production and recent research points more or less unanimously in the direction of what Anderson (2011) concluded from newsroom-based fieldwork in 2008:

“Whereas earlier newsroom sociology emphasized the submersion of audience-centered news judgments in favor of professional codes, the research discussed has documented that the process of ‘deciding what's news’ is increasingly influenced by quantitative audience measurement techniques” (Anderson, 2011: 563).

The influence of these measurement techniques illustrates the way metrics-driven journalism signals a shift in journalism's orientation towards the audience, a shift driven by datafication and one that is a crucial facet of the much deeper transformation in the journalism-audience relationship. It already demonstrates, for instance, the influence metrics have on journalistic role conceptions and how they can lead to an increase in journalists' perceptions of consumer orientation and market considerations in their decision making processes (Hanusch and Tandoc, 2017; Welbers et al., 2016). In extreme cases, audience metrics are even used to evaluate journalists' individual performance through the use of gamified systems that reward them with points and badges (Ferrer-Conill, 2017).

Metrics-informed discussions about what journalism provides and what audiences select have a long tradition in research and are routinely part of public debates on journalism's inability to meet the needs and demands of its audience. However, these discussions can be framed contrarily and can criticize journalism for orientating its attention too heavily towards (click) ratings and audience reach. Reflecting the paradoxical nature of audience ratings, these discussions reveal how, for some, measured audience acceptance is a quali-

ty criterion, while for others, an orientation that relies on ratings represents a danger to quality and professionalism, a warning that what the audience may desire may not actually be in the public interest (Nguyn and Vu; Tandoc and Thomas, 2015).

4 Conclusion: Journalism within a datafied society

The aim of this paper was to situate the datafication of journalism in relation to the datafication of wider society and to describe its present transformation towards a more reflexive, datafied practice that is part of and operating within, a changing media and communicative environment. This process was framed as the transition towards a *datafied journalism within a datafied media environment of a datafied society* and is characterized by the reflexive relationship between journalism and its environment. It also stimulates reflexivity with respect to internal aspects such as, for example, the mutual influence between algorithmically distributed journalism and metrics-driven journalism.

To specify and illustrate this general observation, I have proposed a heuristic made up of four forms of datafied journalism that entangles the stages of news production and consumption with processes of datafication. This entanglement sensitizes us to the most striking transformations in journalism today: data journalism – a new reporting style making sense of the increasing availability of data as a source; algorithmed journalism – emphasizing the particular relevance algorithms have for the dissemination of news items in the online environment; automated journalism, representing the most sophisticated form of datafied journalism; and metrics-driven journalism – highlighting the particular relevance that audience metrics have to the monitoring of news consumption and how it influences the entire news production cycle.

Datafication and the consequences of its presence appear in journalism as if observed under a magnifying glass: Datafied journalism does not only mean that journalism *becomes* datafied; in a much more fundamental sense the four forms of datafied journalism illustrate the reciprocal process between media-related change and social transformation and can best be understood as journalism's response to the datafication of society. In particular, this includes the objects and topics that journalism is supposed to cover, or, put differently, journalism's function in the observation of society: The more the social systems, fields, and domains that journalism is supposed to cover are themselves datafied, that is, the more their social construction relies on data, the more journalism itself needs to be able to make sense of and produce data to fulfil its function for society while managing a need to self-monitor its performance and audience inclusion. This illustrates how journalism is profoundly affected by the processes of datafication, and, by being inextricably intertwined with the development of digital media technologies, is at the same time a main driver of this process:

- Firstly, journalism influences the public discourse on datafication by making it a topic of media coverage; how and to what extent is an empirical question that still needs to be addressed.
- Secondly, journalism increasingly relies on data as a means of observing society and thus possibly sets an incentive for other social domains to produce or make more data available (data PR) as well; a context that needs to be examined in more detail.
- Thirdly, journalism distributes content via various platforms that foster datafication

(e.g. Facebook, Twitter, Google News) and supports the business model of such platforms.

- Fourthly, journalism is increasingly making use of automatically produced content, driving the development of underlying technologies and services.
- Fifthly, journalism drives the increasingly accurate measurement of usage data by its own fixation on metrics.

In sum, journalism has proven itself to be an ideal case study to better understand datafication at a broader societal level. This is because journalism's core functions are becoming increasingly dominated by actors (or actants in the sense of Latour, 2005) that have only recently begun to perform journalism-like services: algorithms built into various platforms and services as well as technology companies developing solutions that produce and simultaneously aim to manage the complexity of the data-generating, data-processing, and data-driven information ecosystem. As such, the four forms of datafied journalism are part of, and a reflection of, broader transformations of the datafied society and public communication. Datafication goes far beyond journalism: it affects the very nature of society's communicative foundations.

Algorithms have not only permeated every stage of the news production process (finding topics, fact checking, selecting, distributing, and producing the news), they are beginning to go beyond these quotidian functions and are beginning to perform journalism-like services independently: they select and filter information, order and structure it, they produce and disseminate content. In doing so they contribute to the construction of the public sphere and social reality. However, algorithms producing news know nothing about the topics they cover and algorithms that disseminate the news know nothing about them either. All news information is transformed into meta data – as information on the amount of news that is used, the networking of this news together with other information: “We are dealing with a situation in which the communication partner is an algorithm that does not understand the content, the meaning, or the interpretations, and works not despite, but because of this” (Esposito, 2017: 254).

All considered, we can also assume that the way news is algorithmically processed and disseminated via social media – without understanding its content – consequently and paradoxically reflects the essential meaning of news in a social sense maybe more appropriately than any idea of broadcasting or publishing: considered from a social perspective, news is about knowing what the others know, not about the content as such (Luhmann, 2000). The crucial point, however, is that the notion of ‘others’ refers to what was previously categorized as ‘the public’. Even if ‘the public’ was never defined in a very precise way, under mass media conditions, it was constructed clearly enough to function as a viable social reference point. In today's transforming media and communication environment we still have to search for a similar, yet alternative, reference point.

5 Acknowledgements

I would like to thank Andreas Hepp, Julius Reimer, Martina Franzen and Cornelius Puschmann for their valuable comments and discussions on this paper. I thank Marc Kushin for his thoughtful proof reading

6 References

- Ananny M (2016) Toward an ethics of algorithms: convening, observation, probability, and timeliness. *Science, Technology, & Human Values* 41(1): 93-117.
- Anderson CW (2011) Between creative and quantified audiences: web metrics and changing patterns of newswork in local US newsrooms. *Journalism* 12(5): 550-566.
- Anderson CW (2013) Towards a sociology of computational and algorithmic journalism. *New Media & Society* 15(7): 1005-1021.
- Ang I (1991) *Desperately Seeking the Audience*. London, New York: Routledge.
- Baecker D (2007a) *Studien zur nächsten Gesellschaft*. Original edition. Frankfurt am Main: Suhrkamp Verlag.
- Baecker D (2007b) Communication with computers, or how next society calls for an understanding of temporal form. *Soziale Systeme* 13(1-2): 409-420.
- Bakshy E, Messing S and Adamic L (2015) Exposure to ideologically diverse news and opinion on Facebook. *Science* 348(6239): 1130-1132.
- Bell E and Owen T (2017) The platform press: how silicon valley reengineered journalism. *Columbia Journalism Review*. Available at: https://www.cjr.org/tow_center_reports/platform_press_how_silicon_valley_reengineered_journalism.php (accessed 30 January 2018).
- Birkner T (2012) *Das Selbstgespräch Der Zeit: Die Geschichte Des Journalismus in Deutschland 1605-1914*. Köln: Herbert von Halem.
- Boczkowski PJ and Anderson CW (eds) (2017) *Remaking the News: Essays on the Future of Journalism Scholarship in the Digital Age*. Cambridge, Massachusetts: MIT Press.
- Borges-Rey E (2017) Towards an epistemology of data journalism in the devolved nations of the United Kingdom: changes and continuities in materiality, performativity and reflexivity. *Journalism*. Epub ahead of print February 2017. DOI:10.1177/1464884917693864.
- boyd d and Crawford K (2012) Critical questions for Big Data. *Information, Communication & Society* 15(5): 662-679.
- Breiter A and Hepp A (2018) Die Komplexität der Datafizierung: Zur Herausforderung, digitale Spuren in ihrem Kontext zu analysieren. In: Katzenbach C, Pentzold C, Adolf M and Taddicken M (eds) *Neue Komplexitäten für Kommunikationsforschung und Medienanalyse: Analytische Zugänge und empirische Studien*. (To be published). Berlin. Berlin: Digital Communication Research.
- Carlson M (2015) The robotic reporter. Automated journalism and the redefinition of labor, compositional forms, and journalistic authority. *Digital Journalism* 3(3): 416-431.
- Castells, M (2000) *The rise of the network society. The information age: economy, society and culture*. Vol. 1. Second edition. Oxford: Blackwell.
- 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. Available at: <http://reutersinstitute.politics.ox.ac.uk/sites/default/files/Editorial%20analytics%20-%20how%20news%20media%20are%20developing%20and%20using%20audience%20data%20and%20metrics.pdf> (accessed 30 January 2018).
- Clerwall C (2014) Enter the robot journalist. Users' perceptions of automated content. *Journalism Practice* 8(5): 519-531.
- Coddington M (2015) Clarifying journalism's quantitative turn. A typology for evaluating data journalism, computational journalism, and computer-assisted reporting. *Digital Journalism* 3(3): 331-348.
- Couldry N and Hepp A (2017) *The Mediated Construction of Reality*. Cambridge: Polity Press.
- Couldry N and Turow J (2014) Big Data, big questions. Advertising, Big Data and the clearance of the public realm: marketers' new approaches to the content subsidy. *International Journal of Communication* 8: 1710-1726.

- Deuze M (2012) *Media Life: We Are Media*. Cambridge, Oxford, Boston: Polity Press.
- DeVito MA (2017) From editors to algorithms. A values-based approach to understanding story selection in the Facebook news feed. *Digital Journalism* 5(6): 753-773.
- Diakopoulos N (2015) Algorithmic accountability. Journalistic investigation of computational power structures. *Digital Journalism* 3(3): 398-415.
- Domingo D, Quandt T, Heinonen A, Paulussen S, Singer JB and Vujnovic M (2008) Participatory journalism practices in the media and beyond. An International comparative of initiatives in online newspapers. *Journalism Practice* 2(3): 326-342.
- Dörr KN (2016) Mapping the field of algorithmic journalism. *Digital Journalism* 4(6): 700-722.
- du Gay, P (ed). (1997) *Production of Culture/Cultures of Production*. London: Sage.
- du Gay, P, Hall, S, Janes, L, Mackay, H, & Negus, K (1997) *Doing cultural studies. The story of the Sony Walkman*. London: Sage.
- Esposito E (2017) Artificial communication? The production of contingency by algorithms. *Zeitschrift Für Soziologie* 46(4): 249-265.
- Ferrer-Conill R (2017) Quantifying journalism? A study on the use of data and gamification to motivate journalists. *Television & New Media* 18(8): 706-720.
- Gillespie T (2014) The relevance of algorithms. In: Gillespie T, Boczkowski PJ and Foot KA (eds) *Media Technologies. Essays on Communication, Materiality, and Society*. Cambridge, Massachusetts: MIT Press, pp.167-194.
- Görke A and Scholl A (2006) Niklas Luhmann's theory of social systems and journalism research. *Journalism Studies* 7(4): 644-655.
- Graefe A (2016) *Guide to Automated Journalism*. New York: Tow Center for Digital Journalism/Columbia Journalism School. Available at: <http://towcenter.org/research/guide-to-automated-journalism/> (accessed 30 January 2018).
- Hanusch F and Tandoc EC (2017) Comments, analytics, and social media: the impact of audience feedback on journalists' market orientation. *Journalism*. Epub ahead of print 21 July 2017. DOI: 10.1177/1464884917720305.
- Hepp A and "Communicative Figurations" research network (2017) *Transforming Communications. Media-Related Changes in Times of Deep Mediatization*. 16. Communicative Figurations Working Paper. Bremen: University of Bremen, Zentrum für Medien-, Kommunikations- und Informationsforschung. Available at: http://www.kommunikative-figurationen.de/fileadmin/redak_kofi/Arbeitspapiere/CoFi_EWP_Hepp-Research-Network.pdf (accessed 30 January 2018).
- Hepp A, Loosen W, Hasebrink U and Reichertz J (2017). Konstruktivismus in Der Kommunikationswissenschaft. Über Die Notwendigkeit Einer (Erneuten) Debatte. *Medien & Kommunikationswissenschaft* 65(2): 181-206.
- Kitchin R (2014) Big data, new epistemologies and paradigm shifts. *Big Data & Society* 1(1): n.p.
- Knight M (2015) Data journalism in the UK: a preliminary analysis of form and content. *Journal of Media Practice* 16(1): 55-72.
- Knorr Cetina K and Reichmann W (2015) Living data in financial markets. Concepts and consequences. In: Süssenguth, F (ed) *Die Gesellschaft der Daten. Über die digitale Transformation der sozialen Ordnung [The society of data. On the digital transformation of social order]*. Bielefeld: Transcript, pp-147-172.
- Latour B (2005) *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford, New York: Oxford University Press.
- Lewis SC and Westlund O (2015) Big Data and journalism. Epistemology, expertise, economics, and ethics. *Digital Journalism* 3(3): 447-466.
- Lindén C-G (2017) Algorithms for journalism: the future of news work. *The Journal of Media Innovations* 4(1): 60-76.

LOOSEN: FOUR FORMS OF DATAFIED JOURNALISM

- Löffelholz M and Rothenberger L (eds) (2016) *Handbuch Journalismustheorien*. First edition. Wiesbaden: Springer VS.
- Loosen W (2015) The notion of the 'blurring boundaries'. Journalism as a (de-)differentiated phenomenon. *Digital Journalism* 3(1): 68-84.
- Loosen W and Schmidt J-H (2012) (Re-)Discovering the audience. The relationship between journalism and audience in networked digital media. *Information, Communication & Society* 15(6): 867-887.
- Loosen W and Scholl A (2017) Journalismus und (algorithmische) Wirklichkeitskonstruktion. Epistemologische Beobachtungen. *Medien & Kommunikationswissenschaft*, 65(2): 348-366.
- Loosen W, Reimer J and De Silva-Schmidt F (2017) Data-driven reporting: An on-going (r)evolution? An analysis of projects nominated for the data journalism awards 2013-2016. *Journalism*, Epub ahead of print 12 October. DOI: 10.1177/1464884917735691.
- Luhmann N (2000) *The Reality of the Mass Media*. Stanford: Stanford University Press.
- Luhmann N (2012) *Theory of Society*. First German edition 1997. Stanford: Stanford University Press.
- Madsen AK, Flyverbom M, Hilbert M and Ruppert E (2016) Big Data: issues for an international political sociology of data practices. *International Political Sociology* 10(3): 275-296.
- Mayer-Schönberger V and Cukier K (2013) *Big Data: A Revolution That Will Transform How We Live, Work and Think*. London: Murray.
- Montal T and Reich Z (2017) I, robot. You, journalist. Who is the author? Authorship, bylines and full disclosure in automated journalism. *Digital Journalism* 5(7): 829-849.
- Münch, R (2002) *Dynamik der Kommunikationsgesellschaft*. Frankfurt a. M.: Suhrkamp.
- Napoli PM (2010) *Audience Evolution. New Technologies and the Transformation of Media Audiences*. New York: Columbia University Press.
- Napoli PM (2014) Automated media: an institutional theory perspective on algorithmic media production and consumption. *Communication Theory* 24 (3): 340-360.
- Nelson JL and Webster JG (2016). Audience currencies in the age of Big Data. *International Journal on Media Management* 18(1): 9-24.
- Nguyen, A and Vu, HT (2018) Numbers behind the news: audience metrics and the changing nature of gatekeeping. In: Nguyen A (ed) *News, Numbers and Public Opinion in a Data-Driven World*. New York: Bloomsbury, pp.113-131.
- Parasie S (2015) Data-driven revelation? Epistemological tensions in investigative journalism in the age of 'big Data.' *Digital Journalism* 3(3): 364-380.
- Parasie S and Dagiral E (2013) Data-driven journalism and the public good: 'computer-assisted-reporters' and 'programmer-Journalists' in Chicago. *New Media & Society* 15(6): 853-871.
- Passoth, J-H, Sutter, T and Wehner, J. (2014) The quantified listener: Reshaping providers and audiences with calculated measurement. In: Hepp, A and Krotz, F (eds), *Mediatized Worlds*. London: Palgrave, pp. 271-287.
- Rogers R (2015) *Digital Methods*. Cambridge, Massachusetts: MIT Press.
- Royal C and Blasingame D (2015) Data journalism: an explication. *#ISOJ: The Official Research Journal of the International Symposium for Online Journalism* 5(1): 24-46.
- Schäfer MT and Van Es K (eds) (2017) *The Datafied Society: Studying Culture through Data*. Amsterdam: Amsterdam University Press.
- Steensen S and Ahva L (2015) Theories of journalism in a digital age. An exploration and introduction. *Journalism Practice* 9(1): 1-18.
- Süssenguth F (2015) *Die Gesellschaft der Daten: Über die digitale Transformation der sozialen Ordnung*. Bielefeld: Transcript Verlag.
- Tabary C, Provost A-M and Trottier A (2016) Data journalism's actors, practices and skills: a case study from Quebec. *Journalism: Theory, Practice, and Criticism* 17(1): 66-84.

- Tandoc EC (2014) Journalism is twerking? How web analytics is changing the process of gatekeeping. *New Media & Society* 16(4): 559-575.
- Tandoc EC and Oh S-K (2017) Small departures, big continuities? Norms, values, and routines in The Guardian's Big Data journalism. *Journalism Studies* 18(8): 997-1015.
- Tandoc EC and Thomas RJ (2015) The ethics of web analytics. Implications of using audience metrics in news construction. *Digital Journalism* 3(2): 243-258.
- Uricchio W (2017) Data, culture and the ambivalence of algorithms. In: Schäfer MT and Van Es K (eds) *The Datafied Society. Studying Culture through Data*. Amsterdam: Amsterdam University Press, pp. 125-138.
- Van Dalen A (2012) The algorithms behind the headlines. How machine-written news redefines the core skills of human journalists. *Journalism Practice* 6(5-6): 648-658.
- Van Dijck J (2014) Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society* 12(2): 197-208.
- Wang Q (2017) Dimensional field theory. The adoption of audience metrics in the journalistic field and cross-field influences. *Digital Journalism*. Epub ahead of print 1 November 2017. DOI: 10.1080/21670811.2017.1397526.
- Welbers K, Van Atteveldt W, Kleinnijenhuis J, Ruigrok N and Schaper J (2016) News selection criteria in the digital age: professional norms versus online audience metrics. *Journalism* 17(8): 1037-1053.
- Williamson B (2017) *Big Data in Education: The Digital Future of Learning, Policy and Practice*. Los Angeles: Sage.
- Yeung, Karen (2017) 'Hypernudge': Big Data as a mode of regulation by design. *Information Communication & Society* 20(1): 118-136.
- Young ML, Hermida A and Fulda J (2018) What makes for great data journalism? A content analysis of Data Journalism Awards finalists 2012-2015. *Journalism Practice* 12(1): 115-135.
- Zuiderveen Borgesius FJ, Trilling D, Möller J, Bodó B, De Vreese C and Helberger N (2016) Should we worry about filter bubbles? *Internet Policy Review* 5(1): 1-16.