News devices

Bounegru, Liliana

university of groningen

GHENT UNIVERSITY

News devices

How digital objects participate in news and research

PhD thesis

to obtain the degree of PhD of the
University of Groningen
on the authority of the
Rector Magnificus Prof. E. Sterken
and in accordance with
the decision by the College of Deans

and

to obtain the degree of PhD of
Ghent University
on the authority of the
Rector Prof. R. Van de Walle
and in accordance with
the decision by the Faculty Doctoral Commission.

Double PhD degree

by

Liliana Bounegru

Supervisors
Prof. Dr. M. J. Broersma
Prof. Dr. K. Raeymaeckers

Assessment Committee
Prof. Dr. M.J.P. Deuze
Prof. Dr. G. Jacobs
Prof. Dr. R.A. Rogers
Prof. Dr. T.A.C. Witschge

Acknowledgements

Many thanks are due to my supervisors, Marcel Broersma and Karin Raeymaeckers, for their continued support, feedback, advice and encouragement throughout the process of developing and writing this thesis, and to the Universities of Groningen and Ghent for providing an institutional home and financial support for this PhD.

I am also most grateful to Tommaso Venturini, who took on the role of informal supervisor and mentor and provided invaluable guidance and advice on all key aspects of this research as well as on the writing process (channeling an inimitable blend of Bruno Latour, Umberto Eco and Howard Becker in his advice on telling about collective life).

The Digital Methods Initiative at the University of Amsterdam, which I consider my intellectual home, has been a continuous source of inspiration and support throughout these years. I would like to thank all of its members: Richard Rogers, Sabine Niederer, Esther Weltevrede, Erik Borra, Anne Helmond, Bernhard Rieder, Michael Stevenson, Carolin Gerlitz, Lonneke van der Velden, Marc Tuters, Natalia Sánchez Querubín, Simeona Petkova, Nadia Dresscher-Lambertus, Jonathan Gray, Saskia Kok, Emile den Tex, Marloes Geboers, Fernando van der Vlist and Stefania Milan. I would particularly like to register my gratitude to Esther Weltevrede and Jonathan Gray for taking time to read various parts of this dissertation and for their useful comments and encouragement; to Richard Rogers, Carolin Gerlitz and Natalia Sánchez Querubín for taking time to meet and discuss this project with me and for their inspiring ideas and suggestions; and to Erik Borra and Emile den Tex for their support with developing the software tools that I used in this research.

The visiting research period at the médialab at Sciences Po, Paris, facilitated by Tommaso Venturini, provided a welcoming environment to develop this work and I would like to thank Bruno Latour, Tommaso Venturini, Mathieu Jacomy,

Christophe Leclercq, Guillaume Plique, Thomas Tari, Vincent-Antonin Lépinay, Barbara Bender, Diégo Antolinos-Basso, Paul Girard, Kari De Pryck, Benjamin Ooghe-Tabanou, Donato Ricci, Audrey Baneyx and Gabriel Varela for lively discussions and feedback during writing workshops, meetings and conversations. I also greatly benefitted from an invitation to visit and present work at the Tow Center for Digital Journalism at Columbia University, as well as subsequent collaborations which emerged from this visit.

The Public Data Lab, of which I am a proud co-founder and member, has provided an ideal environment for the final years of my PhD as well as providing a space for methodological experimentation around the *Field Guide to 'Fake News'*. Thank you for making the journey rewarding: Tobias Blanke, Carolin Gerlitz, Jonathan Gray, Mathieu Jacomy, Lucy Kimbell, Anders Koed Madsen, Noortje Marres, Michele Mauri, Anders Munk, Sabine Niederer, Kari de Pryck, Richard Rogers and Tommaso Venturini.

I would also like to acknowledge co-authors of articles, chapters and reports, collaborators and participants in summer school and data sprints for time, experimentation and reflection which enriched various chapters. I received invaluable feedback from participants at various conference where different parts of this research were presented, including in Amsterdam, Barcelona, Bath, Berkeley, Cambridge, Coventry, London, New York, Oxford, Paris, Perugia, Sheffield, Siegen, Stockholm, Utrecht and Zurich. Anonymous reviewers from *Digital Journalism* and *New Media & Society* provided comments which helped to tighten texts associated with this project. I would also like to note the support of friends and colleagues who supported this work with scripts, tips, calls, conversations and coffees: Sam Leon, Noortje Marres, Craig Silverman, Claire Wardle, Sarah Van Leuven, Stefan Baack, C.W. Anderson, Sylvain Parasie, Michael Stevenson and Phil Howard.

Finally, I would like to thank my family and friends for their love and support, including Lenuta, Carol, Laura, Ruxandra and George, and particularly to my

partner Jonathan who was involved in the day to day joys and challenges of this project.

Table of Contents

1. Introduction

This thesis examines how digital objects participate in news work and how they can be configured for research on digital journalism and its relations with other fields. To this end, it proposes an orientation towards the *news device* as a research topic and approach. I use the notion of news device to refer not only to the digital objects involved in news work but also to a particular way of approaching their participation in these processes. The news device approach calls attention to the ways in which practices and relations are co-articulated with digital objects involved in news work. It also attends to how such digital devices may afford modes of studying these practices. To make the case for the news device approach, I examine the participation of three types of devices in three aspects of news work: (1) the role of the network graph in storytelling, (2) the role of the online platform in journalism coding, and (3) the role of the web tracker in audience commodification.

In developing this approach, I draw on and contribute to several bodies of work which are inspired by science and technology studies (STS): the "material turn" in journalism research (Boczkowski, 2015), and STS-inspired digital media and social research. Overall, the aim is for this approach to contribute to the dialogue between these areas with fresh analytical perspectives and new empirical material.

1.1 Why Study Digital News Work

In Western democracies news work and digital media are today "deeply implicated in one another's existence", as Suchman (2014, p. 129) puts it about media and technology. News work is increasingly taking place in and through a variety of intersecting digital devices, from websites, to search engines, online platforms, apps, bots, web analytics, data analysis and visualisation tools. This is what Lewis and Westlund (2015) call "cross-media news work". Online platforms such as Twitter, Facebook and Instagram, for example, have become sites of experimentation with various journalistic arrangements, from business models, to genres of content, and relations with audiences, sources and other

stakeholders. The implications of these ever-shifting configurations are not yet fully understood, and they are constantly being probed by media organisations, platforms and academic researchers alike. To further complicate matters, it is not only news work that increasingly takes place in and through digital devices, but also our lives, societies, cultures and research practices (Bucher, 2018; Deuze, 2012; Mackenzie, Mills, Sharples, Fuller, & Goffey, 2015; Marres, 2017a; Ruppert, Law, & Savage, 2013). Indeed, online platforms are increasingly understood as "perpetual experiment engines" (Crawford, 2014) that approach society in a "beta-testing" mode (Marres, 2017a, 2017b).[1]

Why does the proliferation of digital objects in news work matter? The reweaving of news work through digital media and technologies has frequently been met with both excitement and concern with regard to its implications for many aspects of news business and practice. For example, there has been a lot of interest in the potential and limitations of these technologies to expand the production and reach of news through experimentation with citizen and participatory journalism (Borger, Hoof, Meijer, & Sanders, 2013; Bruns, 2008; Karlsson, Bergström, Clerwall, & Fast, 2015; Singer et al., 2011) and alternative media (Atton, 2015; Couldry & Curran, 2003; Fenton & Barassi, 2011; Poell & Borra, 2012). There has also been interest in their capacities to improve story discovery and diversify storytelling formats (Broersma & Graham, 2013; Pavlik, 2001; Pavlik & Pavlik, 2017; Thorsen & Jackson, 2018). But there have also been economic concerns associated with digital transformations in the news media industry, the disruption of established business models and the capacities of new models to sustain the journalistic profession (Nielsen, 2016). The advent of digital media has led to questions around the viability of journalism as a profession as we witness a decline in the number of journalists employed by newspapers and their working conditions (Pickard, 2011; Starr, 2012). Concerns have also emerged around legitimation, authority and trust in the relationship between news and its publics, the changing patterns in news consumption and the fragmentation of the audience itself (Blumler, 2010; McChesney, 2003; Peters & Broersma, 2016; Siles & Boczkowski, 2012; Usher,

1 For a discussion of "variability" as new media feature, referring to the potentially infinite versions that digital objects can have, see Manovich (2001). For a discussion of the "perpetual beta" as an approach to developing Web 2.0 or social media products, see O'Reilly (2007).

2018).

Perhaps the most recent wave of unease relates to the scandals about "fake news" (Tandoc, Lim, & Ling, 2018), information disorder (Wardle & Derakhshan, 2017), misinformation, propaganda and political manipulation (Woolley & Howard, 2018) of the past few years. These scandals have further problematised the relationship between the news industry, online platforms and digital technologies and cultures more generally, including around issues of manipulation of public opinion and growing distrust in journalists and experts (McNair, 2017; Tandoc, Jenkins, & Craft, 2018).

1.2 Attending to Digital News Work Through News Devices

These kinds of societal debates are also reflected in the growing body of academic literature about digital journalism. Attending to the complexities of the interactions between digital devices and news industry practices, has led journalism scholars to explore multi-disciplinary outlooks, drawing on fields such as political science, sociology, history, linguistics and cultural analysis (for a review, see Ahva & Steensen, 2017). As I examine further in the next chapter, in recent years there has been growing interest in socio-material perspectives (Ahva & Steensen, 2017) and a so-called material turn in journalism studies (Boczkowski, 2015). These have arisen to address transformations at the interface between news work and digital devices. Coming to the study of journalism as an STS-inspired digital media researcher, I take this body of work as one of the starting points of my project.

Drawing on foundational texts and authors in STS, and particularly Actor-Network-Theory (ANT) and the work of Bruno Latour, these approaches strive to avoid forms of either technological or social determinism, that are said to be present in journalism, media and communication studies literature (Lievrouw, 2014; de Maeyer, 2016; Primo & Zago, 2015). They achieve this through close empirical examination of the mutual shaping or co-production

(Jasanoff, 2004) of digital media and news practices in the situations where they interact. In these accounts both digital objects and news work are considered as socio-material constructions (Boczkowski, 2004a). Digital objects are foregrounded to the extent that they make a difference to news work, and the focus is on when, where and how they make a difference (Kreiss, 2015).

Hence, rather than focusing on large scale changes that digitisation may or may not bring to news work (for better or for worse), in this dissertation I examine how digital news practices and relations can be traced by attending to digital objects in situated contexts of use. In particular, I look at three important but under-noticed objects in digital journalism: network graphs, the GitHub platform, and web trackers; and three aspects of news work: making narrative, making infrastructure and making audience (Table 1). News devices are construed as digital objects situated in particular news work settings. The aim is to explore how each of these devices is implicated in news work and with what consequences.

Area of news work	Making narrative and storytelling	Making infrastructure and coding	Making audience
Digital object	The network graph	The coding platform	The web tracker
News device	The network as storytelling device	GitHub as connective coding device	The tracker as audience marketplace device

Table 1: Types of news work, digital objects and news devices examined in this dissertation

I start with an aspect of news which is perhaps the most familiar and accessible: journalistic representations or news stories. In the rest of the chapters, following the advice that Becker (1982) gives in the context of his work on the sociology of art, I move from *news products* to *news work*, *networks* and "*worlds"* as a way to explore the coordinated activities through which news texts are brought into being. The aim is not to give a definitive and final answer to the question of whether digital transformations of news work are

good or bad and whether we should be excited or fearful of them, but rather to develop means by which questions around the impacts of such transformations in specific settings can be formulated and explored.

The digital objects I selected differ in their kind, provenance, socio-material organisation and functions. As I further discuss in the next chapter, my interest in the digital objects I am studying is informed by issues raised in journalism studies literature and my own professional and research experience. Networks are a common tool in STS-inspired digital social and media research (Moats & Borra, 2018) and I have often used them in my own research work.[2] At the same time I have encountered ambivalence about network analysis and visualisation in conversations with journalists in my professional work for the European Journalism Centre, a non-profit specialising in journalism education and training. The claim about the ontological multiplicity of objects, i.e. the fact that objects may be multiple things all once due to the different ways in which they are enacted in various sites, situations and practices (Mol, 1999; see also Bucher, 2018 on the multiplicity of digital media algorithms) led me to develop an interest in examining how networks are done in the context of journalism, compared to the way they have been put to work in my own field, digital media and social research. Some of the "blind spots" in the study of journalism identified by journalism materiality researchers fuelled my interest in the other two objects. By taking the code sharing platform GitHub as a site of study, I aimed to engage with the critique of the monopoly of the newsroom as a site of study and of news making (Anderson, 2011b; Boczkowski, 2015). I also took up the suggestion to complement more extensive treatments of the editorial sides of news work with the less addressed business sides of news by examining the commodification of news website users through web tracking infrastructures.

Methodologically, the most acclaimed work that takes a socio-material disciplinary approach in journalism research has relied on well-established ethnographic approaches, observation of the work and actions of people and

[2] See, e.g., Bounegru, Gray, Venturini, & Mauri (2018), Venturini, Bounegru, Jacomy, & Gray, (2017) and Venturini, Jacomy, Bounegru, & Gray (2018).

objects, interviews and document analysis (Boczkowski, 2015; de Maeyer, 2016). However digital media and devices are increasingly reweaving not just journalism but also ways of studying and knowing news and other aspects of collective life. The potential of digital traceability is beginning to be acknowledged by STS-inspired journalism researchers (Lewis et al., 2013; de Maeyer, 2016).

This argument that digital devices can be configured as part of the "research apparatus" (Marres & Gerlitz, 2015) is more extensively developed in the context of STS-inspired digital social and media research, which is the second research area that I draw on, and one of my disciplinary homes. The device-centred research approaches that have emerged in these areas draw attention not only to how digital devices shape *practices and relations* (whether social, cultural, economic, political or otherwise) but also how they come with particular ways of *knowing and valuing these practices and relations* (Marres, 2017a; Rogers, 2013; Ruppert et al., 2013; Weltevrede, 2016). In the context of journalism, the latter would include the proliferation of impact and engagement metrics to assess the value of journalism work (Petre, 2015), as well as the use of data from online platforms to study journalism (see, e.g., Bruns, 2016). Device-centred research approaches draw attention to digital devices not just as objects of research but also as possible resources to support research, if configured in the right ways.

The question of the configurability of digital devices into research devices is addressed more extensively in the next chapter. By orienting myself towards these approaches I do not aim to suggest that these should replace existing methods at work in the study of journalism nor that they provide a straightforward way for journalism researchers to integrate data from digital devices in their work. On the contrary, the STS-inspired approaches that I draw on often raise questions about digital traces as simple descriptions of social phenomena and aim to both investigate and interrogate the ways in which digital devices configure action, and the way they process and give access to digital traces. In contrast to understandings of the digital as a straightforward source of data about social phenomena, my aim is to

understand how, to what extent and under what conditions news devices might be successfully enlisted as research devices in the study of news by exploring how they may participate in formulating, configuring and investigating the research object or problem under study.

It is these two propositions, that news work is co-produced through interactions with digital devices and that the latter can be configured into journalism research devices, that I take up in the notion of news device. The aim of this dissertation is to put to test the capacities of this approach to account for the participation of digital media in news work and their implications, through a series of case studies that take digital devices as both an object of study and as part of the research method. This observation brings about another distinction that needs to be made in my objects of study: that between what Rogers (2013) would consider "pre-digital" objects that have seen a renaissance in the context of digital media and culture (the network diagram), and devices and objects that are considered to be "native" to the web, such as the code sharing platform and web tracking devices. I draw upon this distinction because, as I will discuss later in the Introduction, it has implications for the way I methodologically treat these objects.

A final comment to contextualise my approach before going on to the outline of my thesis: both the devices and practices I study and the research areas I draw on are very dynamic and a lot has changed since I started this work. GitHub, the code sharing platform that is the object of study of Chapter 4, has been acquired by Microsoft last year (in 2018). The implications of this acquisition for the platform's business models, uses cases and design remain to be seen. Nevertheless, the acquisition has opened up the space for other platforms to compete with GitHub over developers unsatisfied with the company's move (GitLab, 2018), even though seemingly without success (Finley, 2018; Asay, 2018). Web tracking, the topic of Chapter 5, has also seen major new rules being enforced through the EU General Data Protection Regulation (GDPR).[3] These regulations are aimed at reconfiguring the relationship between website users and the data collector and processor side of

[3] https://eur-lex.europa.eu/eli/reg/2016/679/oj

the advertising market by providing users more awareness and control over the data they share with third-parties while visiting websites online. These rules might also impact the configuration of news audience marketplaces.

A lot has also changed in the journalism socio-materiality research landscape since I started planning work on this dissertation in 2014. Socio-material approaches appear to be on the rise in digital journalism research (Ahva & Steensen, 2017). Work in this area has accelerated and publications have proliferated to the extent that, as mentioned earlier, there is talk of a material turn in journalism research. An entire special issue called "Objects of Journalism and the News" has been dedicated to exploring an "'object-oriented' approach to journalism studies" (Anderson & de Maeyer, 2015, p. 3). Another special issue dedicated to theorising digital journalism has seen an important place offered to socio-technical approaches to understanding journalism (Steensen & Ahva, 2015). The recent *SAGE Handbook of Digital Journalism* (Witschge, Anderson, Domingo, & Hermida, 2016) includes multiple chapters dedicated to socio-technical approaches to understanding the news.

At the same time in my own area, new media and digital culture, a lot has changed in the way of digital research methods. This thesis began in the middle of a moment of accelerating platformisation of the web and associated API-enabled research methods, which I have drawn upon in some of this work. Subsequently, platform responses to misinformation and the recent Cambridge Analytica scandal, have prompted debate about a so-called crisis of API-based social media research (Rogers, 2018), as platforms are increasingly closing their APIs, and as social data is increasingly seen as compromised (see also Bruns, 2018; Venturini & Rogers, forthcoming). These developments make debates about the broader contexts and implications of digital research methods even more pressing, and in the Conclusion I will further discuss some of these implications as well as future research directions.

1.3 Thesis Outline

This thesis starts from the recognition that digital media are shaping not only news work but also ways to create knowledge about news practices. This recognition prompts us to consider how the socio-technical research approaches at work in journalism studies may be combined with device-centred analytical approaches that have emerged in recent digital sociology and new media research.

The central argument of this dissertation is that STS-inspired approaches to the materiality of journalism and device-centred approaches from digital media and social research, in combination, offer modes of studying the participation of digital objects in journalism that can leverage the affordances of digital devices for research.[4] In this dissertation I refer to these as news device approaches and illustrate how the interactions between different digital objects and different aspects of news work can be studied. To summarise, several assumptions underlie this project. The first is that news work increasingly takes places in and through interconnected and competing digital devices, or what Ruppert et al. (2013) call "fields of devices" (p. 19). The second is that news work and digital devices are co-produced though their interactions. Digital devices are not neutral intermediaries but shape the conditions of possibility of news work. News practices do not just make use of digital devices but shape how a device is done and what it becomes in a particular situation. For example, think of the way in which the use of YouTube for opinion manipulation has made us understand the platform as a "misinformation machine" (Tufekci, 2017). Here I align myself with Deuze and Witschge (2018) who call for journalism to be approached as a dynamic, unstable and ongoing configuration of relations, as well as with van Dijck (2013) who proposes digital devices such as online platforms to be understood as "a set of relations that constantly need to be performed" (p. 26). What is important here is not just the recognition *that* the material participates in news making (this is

[4] Weltevrede (2016) defines the "research affordances" of digital media as the research possibilities that emerge from the interactions between researchers and their questions and objectives on the one hand, and the analytical capacities of digital media on the other.

something that has been recognised for a long time) but, rather, how to explore *when* and *how* it participates and with *what consequences* (and to elaborate how it makes a difference, if at all). This recognition informs the overall research question of this thesis: *How can a news device approach help to account for how digital objects participate in and format news work and ways of studying it?*

As mentioned above, this question is addressed empirically by examining the interactions between three areas of news practice and three digital objects. The digital objects are entry points for examining three areas of the digital that are playing a significant role in shaping digital news work: digital visual culture, online platforms, and the online advertising and marketing industries. Methodologically, I use a variety of methods that are sensitive to how digital objects afford, organise and constrain the practices they interact with, from multimodal analysis, to digital methods for single and cross-platform analysis, and visual network exploration.

Through this work I make a number of contributions. Firstly, I empirically examine the role of three less studied but important digital objects in journalism. Secondly, I aim to contribute to a more substantive dialogue between STS-inspired digital journalism research and STS-inspired digital social and media research, by illustrating how methods and sites from digital social and media research can be brought to bear on journalism. Thirdly, I reflect on what this change of sites and methods brings to our understandings of news research and practice, of how digital objects are making a difference in news work and in research, as well as in interactions with other domains.

The thesis is organised as follows. In the introductory chapter, I outline my research project and make the case for why examining the interactions between digital objects and news work is important. I introduce the news device approach to argue that attending to the interactions between digital objects and news from a device perspective may offer fresh angles on digital transformations in particular settings.

Chapter 2 more precisely situates and positions this research project in the

context of STS-inspired journalism studies and within device-centred approaches to digital social and media research. I focus on these bodies of work because of their widely recognised, rich and nuanced approaches to the mutual shaping of technologies and practices, drawing on STS. I discuss the conceptual and methodological aspects of these bodies of work which I bring together in the news device approach. I start by discussing how journalism researchers associated with the material turn approach technical objects as socio-material arrangements, and interactions between news and digital objects as processes of co-articulation. Next, I introduce a related way in which the interactions between digital objects and news work can be approached, namely through device-centred approaches from digital social and media research. These approaches draw on a diversity of fields, from STS, to software and platform studies, and provide a promising way of gaining a rich understanding of digital devices as the material of journalism and as part of the research apparatus for knowing journalism. Finally, the chapter discusses how the news device approach frames the study of interactions between digital objects and news work and describes the development of the three empirical cases.

Chapter 3 is the first empirical chapter. In this chapter I study the narrative affordances of network graphs, the de facto diagram type of the digital age, in the context of journalistic storytelling. The device perspective in this chapter is developed by studying journalistic storytelling or narratives not by taking a particular genre or topic as a starting point, but by taking a key visual and analytic form of the digital age, the network diagram, as a narrative device. The question I ask is: How are network diagrams used for journalistic storytelling and how do network analytics and visualisation shape these stories? To address these questions, I draw on multimodal analysis and graph semiotics to identify the narrative readings that networks elicit in a corpus of journalistic stories. I show that the identified patterns in the narrative readings that network diagrams elicit are shaped by visual network properties and classic network concepts as well as by the journalistic genres in which they are embedded. In doing so I hope to contribute both to the study of journalism and narratives, and to the understanding of the narrative potential of network visualisation, which I discuss as "network stories".

Chapter 4 is a contribution to platform studies, and more specifically to the platformisation of news and to the study of coding in journalism. I address the question of how the online platform GitHub shapes coding practices. I do so by examining how public journalism code development and the making of digital news infrastructures are configured through platform-making activities, which include technical infrastructure configuration and economic imperatives. Drawing on van Dijck (2013), I propose the concept of "connective coding" to capture these processes through which the platform participates in journalism coding to commodify it. I also address the question of how news initiatives participate in the platform by proposing an analytical approach that repurposes connective coding. I examine modulations of practice in a collection of journalism code repositories gathered by Source, a key journalism coding initiative. The focus is on platform-specific forms of coding work and their temporality, and how journalism code is engaged with on GitHub. I find that most journalism code does not keep up with the platform's preoccupation with fresh content and that the platform's culture and dominant publics shape how journalism code is valued.

In Chapter 5 I draw on research into audience economics and media audience markets, and on socio-material approaches to digital media to study audience commodification from the perspective of web tracking devices. More specifically, I take the "fake news" scandal as an occasion to address the relationship between the online advertising and marketing industries and two types of digital cultural production: news and junk news or viral misleading content that mimics the news format. The question that drives this chapter is: How can audience commodification be approached from a digital device perspective? To address this question I extract third-party tracking mechanisms embedded in a collection of pages from mainstream and junk news websites. I examine their audience marketplace configurations from the point of view of the web tracking infrastructures of these sites. This analysis contributes to the study of the business side of news by providing a picture of audience marketplace practices in which legal initiatives such as the recently enforced EU General Data Protection Regulation (GDPR) are aiming to intervene.

In the sixth and final chapter I formulate conclusions about what a news device perspective can bring to the study of news and to journalism research. I summarise the main contributions of this dissertation and reflect on the challenges and implications of this approach for journalism research and sketch a few directions for future research.

2. The “Material Turn” in Journalism Studies Meets Device Perspectives From Digital Social and Media Research

As outlined in the Introduction, this thesis proposes a news device approach to studying the role of digital devices in news and journalism practices. This approach translates difficult but important questions about impact in the relations between news and digital media into a focus on the role of digital objects and devices in enacting[5] various aspects of news work in particular contexts. This outlook combines insights and perspectives from two emerging bodies of research: on the one hand, recent STS-inspired studies of journalism; and, on the other hand, device-centred approaches developed in digital social and media research.

In this chapter I discuss these two bodies of work that I draw on and contribute to. I propose the notion of news device as a way to think through how these approaches can be combined and brought to bear on the study of digital news making, circulation and use. Both areas of work are inspired by approaches from science and technology studies (STS), and particularly from the approach to understanding the making of science and technology developed in actor-network-theory (ANT). STS, as one handbook puts it, "is an interdisciplinary field that investigates the institutions, practices, meanings, and outcomes of science and technology and their multiple entanglements with the worlds people inhabit, their lives, and their values" (Felt, Fouché, Miller, & Smith-Doerr, 2017, p. 1). To understand these areas, ANT proposes an approach that "maps relationships among material entities and artifacts, human actors, and the ideas or symbols associated with them as 'heterogeneous' and open sociotechnical networks" (Lievrouw, 2014, p. 29).

In the next section I introduce recent socio-technical approaches to the study of journalism, on which my approach draws. The aim is not to provide a comprehensive review of this body of work but rather to introduce aspects

5 The notions of "enactment" or "performativity" which are used throughout this thesis are linked to actor-network-theory (see, e.g., Latour, 2005b). As a theory of action (as opposed to a theory of actors), ANT sees socio-technical relationality as not posited between pre-existing entities, but sees these entities as the outcomes or effects of ongoing processes of networked (inter-)action or assembling (Suchman, 2014). According to Law (2004), enactment or performance refer to "the claim that relations, and so realities and representations of realities (or more generally, absences and presences) are being endlessly or chronically brought into being in a continuing process of production and reproduction, and have no status, standing, or reality outside those processes" (p. 159).

that inform the news device perspective: the interest in accounting for the materiality of cross-media news work, the conceptualisation of this work as socio-technical practices, and the identification of areas of news media work in need of further examination. In the second section I introduce device-centred perspectives from digital social and media research and discuss how they address the questions and possibilities opened up by digital media. After introducing these two bodies of work, in the third and final section of this chapter I discuss how the empirical chapters of my thesis were developed, my research design and case study selection.

2.1 The "Material Turn" in the Study of Journalism

2.1.1 Accounting for Materiality by Studying Digital Objects

To develop my approach to the interactions between digital objects and news work, I first turn to socio-material approaches to journalism. The first thing I take from this body of literature is a focus on accounting for the "materiality" of digital journalism. More concretely, this involves accounting for when and how digital technologies, media, objects and their features matter or make a difference in news work situations - from online newspapers (Fortunati, Taipale, & Farinosi, 2015), to interactivity and multimedia (Boczkowski, 2004a, 2004b), blogging technologies (Graves, 2007), content management systems (Anderson & Kreiss, 2013; Rodgers, 2015), application programming interfaces or APIs (Ananny, 2013), Wikipedia infoboxes and cleanup tags (Ford, 2015), hyperlinks (de Maeyer & le Cam, 2015), news site interfaces and commenting technologies (Braun, 2015), software (Usher, 2018), and email and "googleability" (Plesner, 2009).

This is perhaps not a surprising development. As news and journalism are increasingly materialised and saturated by varied first and third-party digital objects, devices and infrastructures, researchers are turning to the study of how these mundane objects participate in news work and with what consequences. Indeed, a notable special issue dedicated to journalism and materiality proposes

to take "objects of journalism" such as those listed above as a starting point for research (Anderson & de Maeyer, 2015). It argues that doing so "provides a new window into the social, material, and cultural context that suffuses our increasingly technologically obsessed world" (Anderson & de Maeyer, 2015, p. 4). Similarly, Boczkowski (2015) and Neff (2015) argue that studying digital objects and their role in news work may have implications for how the domains, actors, places and relations that make up news work are understood and conceptualised.

This growing body of work is increasingly recognised as one of the key modes of theorising and empirically studying the digital transformations of journalism (Ahva & Steensen, 2017; de Maeyer & le Cam, 2015; Lewis & Westlund, 2015; Neff, 2015). Digital transformations of news work are of course more widely discussed by journalism and media researchers outside this small STS-inspired body of work (for reviews of the larger body of work, see, e.g., Domingo, 2006; Mitchelstein & Boczkowski, 2009; Steensen, 2011). But, for me, the most interesting work comes from this smaller area associated with the material turn in journalism research, as it provides a particularly rich understanding of digital objects and technologies.[6]

Socio-material approaches in journalism, media and communication studies have been proposed as alternatives to approaches which are considered to either overestimate or underestimate the capacities of media technologies (de Maeyer, 2016; Lievrouw, 2014; Primo & Zago, 2015). On the one hand, scholars have been accused of assigning media technologies with essentialising effects and responsibility for dramatic changes. On the other hand, they have been blamed for ignoring them, treating them as neutral, or seeing them as mere by-products of interplays of pre-existing social, cultural and economic forces. These are generally known as technological determinist and social

6 Astute readers will point out that there is a longer history of attending to the materiality of media, communication and the news that extends well beyond this recent body of work and which includes well-known names, such as James Carey, Marshall McLuhan and Harold Innis. Instead of focusing on the work of these figures, I will attend to the more recent literature that adopts a "material sensibility" (de Maeyer, 2016). This is because, as we will see, this more recent research has different theoretical underpinnings, which may provide new directions and opportunities for understanding the participation of digital objects in news work.

determinist approaches to news media technologies respectively, even though these "determinisms" in themselves tend to capture the worse misconceptions about both approaches and are in need of reconsideration.[7] In relation to the former, de Maeyer (2016) has argued that:

> The (salutary) move against 'technological determinism' in media and communication studies since the 1960s (Sterne, 2014) has lead *[sic]* to a strong form of social and cultural determinism that, by the 2000s, became the dominant perspective in the field (Lievrouw, 2014: 22). (p. 461)

In relation to socio-cultural perspectives, journalism researchers taking a socio-technical perspective have for some time drawn attention to the fact that it is not just news workers' norms, practices and values that inform the performance of journalism, but also the technologies, infrastructures and material entities mobilised in this work. In relation to media and communication scholarship, Gillespie, Boczkowski, and Foot (2014a), express this point as follows:

> In communication and media scholarship, the overwhelming focus has been on texts, the industry that produces them, and the viewers that consume them; *the materiality of these devices and networks has been consistently overlooked* [emphasis added]. News, in the study of media, has been typically construed as paragraphs on a page, rather than the page itself; the headlines are examined, but not the newsboys who shout them, the teletypes that clatter them out, or the code that now renders them into clickable hyperlinks. (p. 2)

In socio-technical approaches to journalism, materiality is taken to refer broadly to the wide range of digital tools, technologies, artifacts and objects that increasingly participate in news making and often leave traces online (Anderson & de Maeyer, 2015; de Maeyer, 2016). While materiality is a complex notion interpreted differently across various disciplines, one

7 For a re-evaluation of the notion of technological determinism see, e.g., Peters (2011), who argues that the notion is "in desperate need of a critical intellectual history and reappraisal" as "it is a doctrine more often attributed than advocated". See also Peters (2017).

definition of materiality cited in journalism studies literature is that of Leonardi (2012). This definition understands materiality to refer to "the arrangement of an artifact's physical and/or digital materials into particular forms that endure across differences in place and time and are important to users" (Leonardi, 2012, p. 10). In this view the materiality of digital technologies would include the physical features of a technological artifact that matter in particular use contexts. Lievrouw (2014) makes a similar point. Whereas she defines communication technologies as the co-articulation of material artifacts and social arrangements and practices, the materiality of technologies so conceived refers to "the physical character and existence of objects and artifacts that makes them useful and usable for certain purposes under particular conditions" (p. 25).

These media and journalism researchers acknowledge that materiality has been invoked not only in relation to technological artifacts but also in relation to social and institutional practices and cultural objects (see, e.g., Boczkowski, 2015; Kreiss, 2015; Lievrouw, 2014). But the recent calls for materiality in media and news research focus on the materiality of digital artifacts at work in these domains. One of the early examples of this outlook in the context of journalism is Boczkowski's *Digitising the News* (2004a). The book took a different approach to the early literature on digital transformations of newsrooms focused on technological effects, and examined the making of early digital journalism artifacts (such as websites and web pages) by studying multiple situated appropriations of digital technologies and the professional practices, material cultures and local contingencies that shaped early newsroom experiments in online publishing. In a more recent example, Rodgers (2015) studies how a single piece of software, an in-house newsroom content management system, takes on different meanings and plays different roles across different newsroom departments. But, while studying the different appropriations of this software at these different sites, he also emphasises how the software itself shapes or forms the conditions of possibility for news work. This is manifested for example in the way in which audience metrics available through the CMS become increasingly ingrained in editorial thinking and decision making.

The attention paid to the role of objects in enacting news work is accompanied by an understanding of news and journalism as socio-technical practices (Anderson, 2015; Domingo & Wiard, 2016; Lewis & Westlund, 2015). Lewis and Westlund (2015) describe a socio-technical outlook not as a way to bring into focus how technology is impacting journalism but to foreground the ways in which organisational logics, relations and processes of news work are becoming increasingly entangled with a multitude of intersecting technological mediations and devices. The interactions between technology and news practices are understood as processes of "co-production" or "mutual shaping" (Domingo & Wiard, 2016; Gillespie, Boczkowski, & Foot, 2014b; on the notion of co-production, see also Jasanoff, 2004), even though Lievrouw (2014) believes that this remains somewhat of an "unfinished project" (p. 24).

So, while the term materiality or material turn might seem to enact a division between the study of journalism materiality or technology, and that of journalistic culture, practice, texts and history, a closer look at the body of work assembled under these labels suggests that what it in fact aspires to do is to break down these divisions and to problematise technology by examining the historical, social, cultural and political contexts in which it is being produced, transformed, appropriated and used (see, e.g., de Maeyer & le Cam, 2015). In the context of studying spam and online activist groups, Brunton and Coleman (2014) argue that a material sensibility does not involve a return to hardware and technical infrastructure as a way to get to reality. On the contrary, they argue:

> When we peel back that deepest layer of materiality, we find people and practices underneath: populations of users, and the "superusers" who operate close to the metal in their work, including system and net administrators (sys/net admins), hackers, and spammers in complex, contingent, ambiguous relationships (p. 77).

2.1.2 Conceptual and Methodological Considerations

While this body of work takes up different journalistic concerns, from technological innovation more generally (Domingo, Masip, & Costera Meijer, 2015; Micó, Masip, & Domingo, 2013; Schmitz Weiss & Domingo, 2010), to particular objects of journalism (Anderson & de Maeyer, 2015; de Maeyer & le Cam, 2015; Usher, 2018), the attention to socio-materiality that has emerged in journalism studies in the past years is perhaps most prominently associated with science and technology studies (STS) and particularly with actor-network-theory or ANT (Ahva & Steensen, 2017). ANT has been described as holding "the most fundamental implications for our analysis of journalism in the digital age" (de Maeyer, 2016, p. 463). As the name suggests, STS (and its associated ANT) emerged in the context of social studies of natural science and technology, and the socio-material approaches developed in journalism studies rely on these foundational texts (particularly the work of Bruno Latour, but also that of Michel Callon, John Law and others).[8]

Concepts and approaches from STS and particularly from ANT inform research with a socio-materiality outlook not only in journalism but also in media and communication more broadly (for a comprehensive discussion, see Gillespie et al., 2014b). In studies of journalism, ANT is typically mobilised to account for technological actants, their agential capacities, and the way they participate in the doing of journalism (Lewis & Westlund, 2015, 2016; Micó et al., 2013; Primo & Zago, 2015; Plesner, 2009; Schmitz Weiss & Domingo, 2010). The agential capacities of technologies in these studies are not seen as essentialised forces with universalised effects or consequences. Agential capacities of technologies become traceable in situated contexts of action, through associations with other actors. So technical agency is seen as different from human agency and does not refer to a sort of an independent consciousness of technical systems, but rather to the ways in which socio-

[8] While the works of Latour, Callon and Law have prominently featured in the reception of ANT in journalism studies, it should be noted that there are many others who have made seminal contributions to developing and establishing this approach whose works are deserving of more attention in relation to the study of news, including, for example, Madeleine Akrich, Cécile Méadel, Annemarie Mol and Susan Leigh Star.

technical systems facilitate, afford and constrain action.

Finally, while this approach is thriving in journalism research, it also raises methodological issues when it comes to approaching digital news work. While established work in this area is informed by ethnographic fieldwork on offline sites such as the newsroom, observation, interviews and document analysis, there is also an interest in methodological experimentation. Boczkowski (2015) argues that accounting for the participation of digital objects such as algorithms in news making will require "broadening our methodological apparatus" (p. 67). De Maeyer (2016) hints to digital traceability as one way to orient methods for studying materiality: "Digitisation shines a new light on the question of materiality by offering traces of what previously may have been gone unnoticed" (p. 461). While abandoning the newsroom and other offline sites is certainly not a useful way forward, complementing the newsroom with online sites of study, and the methodological apparatus of journalism research with methods informed by the specificities of digital devices, may be a direction worth testing given that news and information flows are increasingly being curated through digital platforms and devices. Moreover, given that the online and the offline are not separate but entangled, the online may become a way to access and characterise other sites and actors (Rogers, 2013).

2.1.3 Some Commitments and "Blind Spots" in Journalism Research

Another aspect developed in the journalism socio-materiality literature that proved useful for my research project is its discussion of commitments at work in existing approaches to journalism and what are considered to be, in the words of Pickard (2017), "blind spots" in journalism research: areas of news work in need of more sustained attention. I discuss some of these below not as a critique but in the spirit of taking up some of these concerns in my empirical work.

Recent studies emphasise the anthropocentric tendencies of journalism research, its newsroom-centrism and the emphasis on editorial work at the

expense of other areas of news work. These concerns are not only present in the journalism socio-materiality literature — on the contrary they are increasingly being raised by researchers outside this area too (see, e.g., Deuze & Witschge, 2018; Pickard, 2014; Zelizer, 2004). But in this section I will focus on discussing them in relation to this body of work that informs the research approach developed in this dissertation.

The anthropocentric tendencies of journalism scholarship refer to the dominant view of journalism practice as the domain of human actors, particularly journalists, and the scholarship's focus on the practices of this limited set of actors (Boczkowski, 2015; Lewis & Westlund, 2015; Primo & Zago, 2015). These authors argue for multiplying the types of actors acknowledged in journalism research, one the one hand towards less studied human actors (including programmers, technicians, graphic designers, project managers, marketers, sales representatives, customer relationship managers and bloggers), and one the other hand towards non-human actors or technological actants. The notion of technological actants refers to material objects that make a difference in the course of action of actors (Latour, 2005b). In the context of journalism research technological actants are described as being "inscribed and instructed by humans, socially constructed to suit journalistic, commercial, and technological purposes within news organizations" (Lewis & Westlund, 2015, p. 24). Examples of these might include: email, CMS-es, APIs, mobile applications, and algorithms. Finally, abandoning anthropocentric in favour of hybrid perspectives in journalism and media more generally, these authors argue, also involves acknowledging that human subjects might not always be at the centre of communication environments and that "there are times and places when and where we are not fully in control of our machinescapes" (Neff, Jordan, McVeigh-Schultz, & Gillespie, 2012, p. 312; see also Anderson, 2016; Primo & Zago, 2015).

Secondly, the newsroom-centrism of journalism scholarship refers to the unquestioned privileging of the newsroom as the central locus of journalism and as its site of study (Anderson, 2011b, 2016; Boczkowski, 2015; Primo & Zago, 2015). These authors argue not that the newsroom is no longer a key

locale of news work but that its place and role need to be empirically tested by examining its relationship to the wider news, media and communication ecosystem.

Finally, these studies often take issue with the way in which news and journalism research is dividing the news world into separate domains of study, or with the way in which attention is distributed across these domains (Boczkowski & Siles, 2014; Domingo et al., 2015; Domingo & Wiard, 2016; Lewis & Westlund, 2015). Lewis and Westlund (2015) draw attention to the unequal distribution of attention across the business, editorial and technological side of newsroom work, where the editorial side receives most attention. They call for a more equal emphasis on all sides and phases of newsroom work, as well as the interconnections between them (Lewis & Westlund, 2015). The business side of news work is said to be in particular need of more sustained attention (Lewis & Westlund, 2015; Pickard, 2017). Going a step further, Boczkowski and Siles (2014) argue for a "cosmopolitan" approach to the study of media, including news media, where the analytical divides between production and consumption on the one hand, and between content and material dimensions on the other would be removed in order to show the interconnections of these dimensions in practice.

2.2 Device Perspectives in Digital Social and Media Research

So far I have discussed the research sensibility towards the materiality of digital journalism that is developed by socio-technical approaches to the study of journalism, and which I draw on in the news device approach. To further elaborate this approach I turn to a number of STS-inspired device-centred perspectives from digital social and media research.[9] These approaches are

9 In this thesis I use "STS-inspired digital social and media research" not to indicate a unified research area but as a shorthand for a set of interesting and varied work that emerges at the intersection between digital social research, new media studies and STS and which informs my understanding of device-centred research approaches. My discussion of it is not meant to provide a comprehensive review but rather to capture the key aspects of device-centred

given different names, including "digital methods" (Rogers, 2013), "interface methods" (Marres & Gerlitz, 2015), "device-driven research" (Weltevrede, 2016; making an analogy with "data-driven" research), "device-aware sociology" (Marres, 2017a), and "device-centred perspectives" (Marres, 2012a). In this thesis I adopt the term "device-centred perspectives" to refer to these approaches collectively. When discussed individually I use each different name.

I draw upon these because they provide rich and nuanced approaches to the interactions between society, knowledge creation and digital media and technologies. The news device approach that I suggest channels these to help enrich the treatment of interactions between news, digital devices and ways of studying them. Drawing on these, it suggests attending both to how the digital offers sites, technologies and practices through which news work can play out, as well as to its possibilities and limits for studying both news and digital devices.

The concept as deployed in these bodies of work has a number of lineages. In STS, and particularly in ANT, the concept of device has been used to introduce objects and equipment deployed in the production of science into sociological inquiries about this topic. These objects are understood as socio-technical arrangements that serve particular purposes.[10] In their seminal *Laboratory Life,* Latour and Woolgar (1986) propose the notion of "inscription devices" to account for the work that items of laboratory equipment,

---

perspectives to interactions between society and digital technologies. For broader discussions of the interplays between media studies, sociology and STS, see, e.g., Badouard, Mabi, Mattozzi, and Schubert (2016), Boczkowski and Lievrouw (2008), Gillespie, Boczkowski, and Foot (2014), Sørensen and Schubert (2015), and Wajcman and Jones (2012).

10 The notion of socio-technical arrangement, used often in this thesis in a number of variations, including heterogenous or socio-material arrangement or assemblage, is a key concept in STS and particularly in ANT. In the words of Callon (2004), "the socio-technical *agencement* is one of the central concepts of the anthropology of the sciences and technologies and, more particularly, of actor-network theory (ANT): describing a combination of human beings and technical devices that are caught in a dynamic configuration (the *agencement* acts), it emphasizes the composite and distributed character of all action and the impossibility of definitively separating humans from technologies. It is *agencements* that are primary and which give their meaning to categories such as States, markets, families" (p. 121). Understood this way, the agential capacities of devices are not the sum of the individual agencies brought together in a device but are the result of encounters between subjects and technical objects. This conception sees entities themselves as effects or outcomes of relationality (Suchman, 2014). Moreover, the notion should be thought of an ongoing "process of assembling rather than a static arrangement" (Bucher, 2012b, p. 40).

individually or in combination, do in the construction of scientific facts. Particularly inspiring for my own notion of news devices has been Muniesa, Millo and Callon's (2007) notion of "market devices". Muniesa et al. argue for including objects – such as pricing models, accounting methods, trading protocols, benchmarking procedures and financial charts – more prominently in empirical studies in economic sociology. Here market devices are not just objects that represent various economic aspects, but are understood as "material and discursive assemblages" that participate in the configuration of markets (Muniesa et al., 2007, p. 3). Similarly, news devices are digital objects of various provenances that participate in the configuration of news work in various ways. The notion aims to draw attention to how digital objects are not neutral carriers but participate in the configuration of news work by helping to render various aspects of collective life into news. For example, Chapter 3 illustrates how network diagrams help to render aspects of collective life into stories. And Chapter 5 shows how web tracking devices help to render news site visitors into audience commodities.

At the same time the notion of devices in STS draws attention to the fact that objects are inscribed with particular logics and assumptions that inform their operation: technology "both embeds and is embedded in social practices, identities, norms, conventions, discourses, instruments, and institutions" (Jasanoff, 2004, p. 3). Indeed, calling these news devices draws attention to how digital objects become implicated in news work but does not mean that they serve exclusively or primarily the purposes of journalism. The digital devices examined in this dissertation often have provenances external to the news domain and thus are inscribed with other logics and purposes. This may pull news practices that deploy them in directions that may not necessarily be entirely favourable to news institutions, as is discussed further in the empirical chapters.[11]

In this literature the notion of device is also associated with Foucault's (1980) "dispositif", often translated as "apparatus". Foucault uses this notion to

[11] For a discussion of the tensions that characterise the relations between news institutions and online platforms between short-term gains in some areas of news work versus longer-term concerns about more systemic platform dependencies, see, e.g., Nielsen & Ganter (2018).

describe shifting systems of relations between heterogenous elements "consisting of discourses, institutions, architectural forms, regulatory decisions, laws, administrative measures, scientific statements, philosophical, moral and philanthropic propositions" (p. 194). The apparatus is set in place in response to a stringent need and serves a particular strategic function, for example to control phenomena such as criminality in the case of the prison system. Heterogeneity in the context of the apparatus is meant to draw attention to its discursive and non-discursive or material composition.

These lineages inform to various degrees the digital device perspectives that I will discuss below and on which my news device approach draws. In what follows I will situate and describe these perspectives, drawing out aspects that inform my empirical work on the interactions between digital objects and news practices.

2.2.1 Digital Devices as Objects of Study

In response to early accounts of the digital that imagined it as a monolithic development with uniform, pre-determined effects, Ruppert et al. (2013) call for approaching the digital as the proliferation of varied digital media, devices and objects in specialised and mundane settings. Device-centred approaches to social research call for treating these digital devices with a sensitivity towards their "double social life" (Law, Ruppert, & Savage, 2011). This means that, on the one hand, we should explore how devices such as social media platforms, mobile applications, blogs, websites, wikis and search engines materialise social, economic, cultural and other relations in particular situations (Mackenzie et al., 2015; Marres, 2017a; Rogers, 2013; Ruppert et al., 2013; Savage, Ruppert, & Law, 2010). One the other hand, we should explore how digital devices themselves are shaped by social worlds (Langlois, McKelvey, Elmer, & Werbin, 2009). These two lines of inquiry are present in understandings of dominant online devices – e.g. Google Web Search, Facebook, Twitter and Wikipedia – as interplays between device-specific materiality, economic imperatives, social practices and cultures of use

(Weltevrede, Helmond, & Gerlitz, 2014).

Often the terms "device", "object" and "medium" are used interchangeably in these approaches. The term "device" refers not just to social media platforms but also to applications or apps, search engines and other digital media and technologies. Moreover, the term object can refer to the various elements – e.g. tags, links, tweets, likes, online profiles, date stamps – that larger socio-technical arrangements such as platforms organize. And there are of course the more mundane uses of the term device to refer to tools, gadgets or tactics.[12] Regardless of the material shape that they take, as Law and Ruppert (2013) remind us, referring to these as "devices" indicates a particular treatment as "patterned teleological arrangements", i.e. as more or less stable arrangements of heterogenous elements that do particular things. For example, the notion of news device is intended to draw attention to how news is done or performed through the interactions between journalistic work and digital objects and how the objects too are shaped through their situations of use and other social, legal and economic aspects. As I will discuss later in this section, the news device approach explores how news devices can become "research devices" by making analytical affordances of digital objects part of the research apparatus (Weltevrede, 2016).

Several materiality-sensitive research areas have emerged to attend to the specificities of digital media. In what follows I will introduce two of these that are particularly relevant for my empirical studies: software studies and platform studies. Both software and platform studies are diverse fields and both largely foreground the underlying material and technical infrastructures of digital devices and their cultural, political and economic implications. Software studies emerged in the context of growing interest in the theorisation of web and social media as spaces for cultural practices, content circulation and user participation. It responded to a concern that software, in spite of its ubiquity in today's societies, "has begun to sink into its taken-for-granted background" (Thrift, 2005, p. 153). Departing from what Bucher (2012) calls "usage studies" and the narrow focus on human activity, software studies investigates "the role

12 https://en.oxforddictionaries.com/definition/device

of software in forming contemporary culture, and cultural, social, and economic forces that are shaping development of software itself" (Manovich, 2008, p. 5). That is, software studies draws attention to the "conditions of possibility that software establishes" (Fuller, 2008, p. 2), i.e. the way in which software and protocols create the conditions of possibility for participation, sociality and knowledge. An example of a contribution to this field is Bucher's (2012) "The Friendship Assemblage: Investigating Programmed Sociality on Facebook" (2012). The article examines not how individuals use Facebook to establish and maintain friendships but rather how friendship is configured on Facebook through software actors such as the "People you may know" algorithm. This algorithm guides users in discovering new friends according to platform logics. Another software actor, the News Feed algorithm, gives more visibility to some friends than others, based on principles of similarity.

The related field of platform studies has its origins in the study of video game platforms and today extends to social media platforms. In the context of profit-driven commercial media systems, it responds to calls to investigate the interplays between the technical infrastructures of platforms and use cultures, or "the connections between platform technologies and creative production" (Bogost & Montfort, 2007, p. 1). Platform studies "highlight how platforms' affordances simultaneously allow and constrain expression, as well as how technical, social, and economic concerns determine platforms' structure, function, and use" (Plantin, Lagoze, Edwards, & Sandvig, 2016, p. 6). The aim is to explore how the material affordances of platform infrastructures shape participation, sociality, cultural production and other activities that are enacted through them. This is done by investigating platform features and objects such as user profiles, ranking algorithms, trending algorithms, social buttons, hashtags, metrics, engagement counters, APIs and so on. Researchers draw attention to programmability as a defining technical characteristic of platforms, which enables ecosystems of apps and services to emerge around them (Bogost & Montfort, 2009; Helmond, 2015a). A notable example of such a study is Gerlitz & Helmond's (2013) "The Like Economy: Social Buttons and the Data Intensive Web". The article examines how Facebook's Like button, related social plugins and the Open Graph, configure not only sociality and

participation in the platform but the very infrastructure and economy of the web, by embedding tracking devices in websites and thus enlisting websites into a "like economy", one where social engagement is of economic value.[13]

Particularly relevant for my empirical work are the concepts of "platformisation of cultural production" and "platformisation of news" that connect investigations of the technical infrastructure of platforms with their economic imperatives in the context of digital cultural production (Nieborg & Poell, 2018; van Dijck, Poell, & de Waal, 2018; see also Helmond, 2015a for the related notion of the "platformization of the web"). The concept draws attention to the increasingly dependence of digital cultural industries on big online platforms collectively known as GAFAM (Google, Apple, Facebook, Amazon, and Microsoft). It calls for empirically studying the mutual shaping of infrastructural, governance and economic mechanisms of online platforms with processes of cultural production and circulation. In such studies social media platforms are not just distribution channels for content, as the web had previously been conceptualised, but rather "articulations of technical, corporate and media logics" (Langlois & Elmer, 2013, p. 2) which seek "to establish the conditions within which content can be produced and shared and where the sphere of agency of users can be defined" (Langlois, McKelvey, et al., 2009, p. 5). Approached this way cultural objects such as YouTube videos are not just content but "thick digital objects" that enable the study of the mutual articulation of participatory culture and platform economic logics (Langlois & Elmer, 2013, p. 12).

In the context of my research work, software studies, platform studies and the platformisation of cultural production are mobilised in Chapter 4 where I examine how the code sharing platform GitHub establishes the conditions of possibility for journalism coding in alignment with its economic imperatives, and in Chapter 5 where I trace the connections between economic aspects of digital cultural production (the production of the audience commodity), and their implications for an important part of the infrastructure of news: the news

13 For more extensive reviews of software and platform studies see, e.g., Bucher, (2012b) and Helmond (2015).

website.

2.2.2 Digital Devices as Resources for Research

A second aspect that I incorporate in the news device approach and my empirical studies from digital social and media research, is the treatment of digital devices as offering modes of knowledge creation. In the previous section, I showed how digital social research, digital culture and platform and software studies scholars have drawn attention to how digital devices are increasingly performative of contemporary social life and cultural production. In this section, I turn to how digital devices also increasingly participate in the analysis of collective life and how they afford new modes of research (Marres, 2017a; Rogers, 2013; Ruppert et al., 2013; Weltevrede, 2016).

Ruppert et al. (2013) argue that "digital devices and the data they generate are both the *material* of social lives and form part of many of the apparatuses for *knowing* those lives" (p. 3), be they sociological or not. In other words, digital devices and platforms can be seen as both an object of study and as a resource for research.

Such assertions are informed by the capacities of online technologies to generate digital traces (Venturini & Latour, 2009) or "transactional data" (Savage & Burrows, 2007) about these enactments and to extensively document them in structured ways. According to Marres (2017a): "what distinguishes the digital technologies of today – what sets them apart from the 'Web' and 'information and communication technologies' (ICT) that went before – is their extensive capabilities for monitoring, analysing and informing social life" ("What is Digital Sociology?", para. 2).

Claims about the potential of digital devices to act as a resource for social research are not exclusive to and neither primarily associated with STS-inspired digital social and media research, even though the notion of "digital methods" (Rogers, 2013), which I introduce below, has gained quite some currency in

recent years. The promise of digital traceability for social research is most prominently articulated in debates about "big data" and computational social science, which explore how digital data can be used to understand human and group behaviour on a large scale (see, e.g., Kitchin, 2014; Lazer et al., 2009). This articulation of the promise of digital device data and computational methods for research has also found its way in methodological discussions in digital journalism research (see, e.g., Boumans & Trilling, 2016; Lewis et al., 2013).

STS-inspired device-centric approaches depart from this more prominent, computationally-inflected orientation towards digital data in social research, which they see as not well equipped to address essential questions raised in their disciplines, such as those regarding the mutual shaping of devices and social worlds. In the words of Marres (2017a), the limitation of this approach is

> the very framing of the object of enquiry: as long as the object of computational social science is defined as human behaviour or experience, it is not well-positioned to address – as a positive, empirical topic for digital sociology – the question of how technology, sociality and knowledge – and much else besides – interact in digital societies. ("Problems With Digital Ways of Knowing Society", para. 2)

Precisely due to the question that Marres (2017a) raises above, big data critiques have argued that digital data is biased in various ways (see, e.g., Tufekci, 2014). Device-centric approaches take questions of digital bias seriously but do not abandon devices as resources for research. Marres (2017a) suggests that digital data bias is a problem to be corrected when online data is used as a mirror or window into human and social behaviour. But for approaches that aim to account for how social action is *co-produced* with devices, how devices shape action and data becomes a vital topic of investigation ("Problems with Digital Ways of Knowing Society", para. 1-4).

As alternatives to big data approaches, device-centred perspectives develop their own ways to configure digital devices as part of the research apparatus. Marres (2017a) formulates this proposition as follows:

> Important social research methods are already built into digital infrastructures, devices and practices, even if they currently tend to serve other-than-sociological ends. I argue that it therefore is our task to test and develop the capacities of these methods-devices for social enquiry, so that they may better serve its purposes. While digital architectures constrain social research in many ways, they are also to an extent configurable: the digital application of method requires a continuous mutual adjustment of research question, data, technique, context and digital setting. ("Introduction", para. 5)

Indeed, methodological experiments under the labels of digital methods (Rogers, 2013) and interface methods (Marres & Gerlitz, 2015) have drawn attention to how "methods embedded in online devices" can be repurposed or configured for social, cultural and media research (Rogers, 2013, p. 1). This is to say that techniques of data capture and analysis inscribed in digital device features and functions which enable social life to be performed online, may be repurposed for particular types of social and media analysis. In other words, the metric-intensive and networked nature of activities and content online, through hyperlinking, likes, tweets, profile categories, tags, etc., holds, as Weltevrede (2016) puts it, particular "research affordances", i.e. it invites and facilitates particular modes of analysis. For example, hyperlinks have been used to study the "politics of association" between actors with hyperlink analysis (Rogers, 2004; Rogers & Marres, 2000), and Facebook page likes have been repurposed to examine associations between pages through "page like network" analytical techniques (Rieder, 2013).

In order to investigate the role and capacities of digital media in society, these proposals distinguish methods specific to each medium ("methods of the medium") from the use of more established social research methods such as the survey (Rogers, 2013). The positive valuation of methods native to the performance of social life with digital devices in this research programme may be seen as linked to a longer tradition in sociology to see methods for documenting and accounting for social life as always already part of everyday life (for a discussion see Marres, 2017a, "Changing Relations Between Technology, Sociality and Knowledge", para. 1-2). Digital media devices are

increasingly important actors that mediate the documentation of social life. The proposal of digital methods in the programme formulated by Rogers (2013) and the Digital Methods Initiative, of which I am a member, is to repurpose the methods by means of which social and cultural life are curated in order to understand *both* social life and the curation process.[14]

According to Weltevrede (2016), digital methods are medium-specific in the sense that they encourage a sensitivity towards the socio-material specificity of each medium, i.e. to the data, methods, objects, practices and cultures through which social life is performed in and with a particular device. Treating devices as interplays between their materiality, human actors and their cultures and practices of use (Weltevrede et al., 2014), has implications for how we draw out their capacities as research devices. When treated this way, repurposing devices for social research does not just mean using them as sources of data. It also means accounting for how these interplays enact the studied phenomenon *and* the knowledge produced about it, and devising ways to make what Weltevrede (2016) calls their "operational capacities" (e.g. the way in which they store, sort, filter and order content), as well as their use cultures, productive for social research. That is to say, the "methods embedded in online devices" – i.e. how devices such as platforms treat objects such as tweets – become at once the object of investigation *and* part of the resources drawn upon to study them (Rogers, 2013).

While the arrival of digital traceability in social sciences and humanities research reopens debates about qualitative and quantitative methods (Venturini & Latour, 2009), a qualification of digital methods along this axis does not capture perhaps the most relevant aspects of this approach. By focusing on digital devices not just as sources of data about the social but also as curating and formatting sociality and ways of knowing it, Weltevrede (2016) describes such device-driven approaches as "thick methods". Such methodological approaches call the researcher to navigate ambiguities and tensions between device provenance, resources and assumptions and the researcher's own goals through inventive configuration (Marres, 2017a; Marres & Gerlitz, 2015;

14 See https://digitalmethods.net

Weltevrede, 2016). Inventive configuration combines the formulation of research questions that allow to surface insights about the mediated issues and the mediating devices, with qualitative and quantitative analysis informed by the analytical affordances of devices, and distant and close reading of digital data sensitive to the specificities, assumptions and cultures of use of the device from which it originates.

A final point is important to be made in relation to device-centric or medium-specific digital methods approaches which I use in my empirical research. I dedicated most of the discussion in this section to explaining and situating the way in which device-centric approaches complicate the relationship between digital devices and methods. This is because this relationship is not always intuitive at first sight and it sometimes gives rise to misconceptions through association with the more dominant approaches of computational social science.

But what makes this a rich research approach is also what makes it challenging. Marres (2017a) cautions that the capacities of digital devices for social research should not be taken for granted. In the case of platforms, following the analytical affordances of the medium, which often offer metrics of frequency of occurrence (e.g. how many times a tweet has been retweeted, a post liked and so on) might pull the researcher towards forms of analysis privileged by the platform (e.g. popularity, influence and trends), and might constrain the research towards particular questions and directions. Issues of data access and data collection may further steer the research direction (for a discussion of how data collection techniques such as scraping and API calling shape research, see Marres & Weltevrede, 2013). This does not mean that digital data should be abandoned. But rather that, through the configuration of the entire research apparatus, the researcher can devise techniques to resist and push against this pull of the device and ensure that questions, device, method and data collection are all aligned. One such technique might be the move from frequency metrics to relational analysis that might reveal associational dynamics over time, such as in the case of co-hashtag analysis (Marres & Gerlitz, 2015; Marres & Weltevrede, 2013). These modifications often appear subtle. An

example of such a modification would be undertaking a qualitative exploration of the nature of public pages that engage with stories on Facebook to resist the pull of platform metrics in the direction of quantitative assessments of engagement, as was the case in another one of my studies (Gray, Bounegru, & Venturini, forthcoming). But such techniques are necessary to enable social and media research with digital devices.[15] For this reason Marres (2017a) recommends that the configurability of devices for social or media inquiry should be tested empirically rather than assumed. The trials to align the affordances of platforms with research concerns will become evident in the empirical chapters of this dissertation.

Another challenge pertains to the types of research that device-centric approaches enable. By taking digital platforms not just as sources of data but as performative of social life and methods, digital methods approaches open up two different directions of research: studying digitally curated social phenomena and studying the mediating devices themselves (Marres & Moats, 2015). The researcher can configure the research apparatus to privilege one direction over the other but this distinction may be difficult to sustain in practice as the research often does not only capture societal dynamics but also the operations of various digital media, their medium specificities and device cultures (Marres & Moats, 2015; Weltevrede, 2016). Given the inseparability of medium and phenomenon in practice this may be seen as a virtue as it opens up questions about the boundaries of research objects across disciplines (Marres & Moats, 2015).

A final risk may be seen in attending to complexities of the empirical world through the device approach. Namely that "the devices under study come to appear as all-powerful" in the configurations that they engender (Birkbak, 2016, p. 26). Moreover, by singling out particular devices for examination, the researcher may participate in stabilising their dominance and legitimating the particular ways of knowing society that they give rise to. Indeed, actor-network theory, on which the device approach draws, has been criticised for

15 For more on tactics to align the device affordances with social research questions and precautions that a social researcher needs to take when using data from digital platforms for social research, see, e.g., Venturini, Bounegru, Gray, & Rogers, 2018.

naturalising a conception of truth based on the strength of alliances that make up an "actor-network" and for being less able to account for dissenting voices that are being silenced and thus do not leave traces (see, e.g., Amsterdamska, 1990; Star, 1990; Venturini & Munk, forthcoming). However, the description of authoritative devices does not need and indeed does not follow the narrative offered by the actors themselves as the researcher aligns the research apparatus with questions informed by social and media research sensitivities (for a discussion, see also Birkbak, 2016).

Finally, as the web changes through processes such as platformisation (Helmond, 2015a), so do its methods of study. While early digital methods repurposed digital objects such as hyperlinks to examine associations between websites, today digital methods explore the research affordances inscribed in single platforms through objects such as retweets, likes and hashtags and in objects that travel across platforms such as web tracking devices (Rogers, 2017). The Digital Methods Initiative offers dozens of device specific tools, each of which incorporates a tactic to make analytical affordances of a device productive for social and media analysis. For the purposes of this research I use some of these tools and, when needed, co-developed new tools[16] and perspectives with the Initiative to approach devices that the Initiative has not previously explored (such as GitHub).[17]

2.2.3 Towards a News Device Approach

I conclude this section with a few final reflections on what might be involved in bringing device-oriented perspectives from digital social and media research to bear on the study of news work. When applying these to the study of news work practices, I refer to them as the news device approach, to draw attention to how they participate in and attend specifically to the study of digital news

16 For the set of tools developed with the Digital Methods Initiative in the context of this project to extract and analyse data from the GitHub API, see https://wiki.digitalmethods.net/Dmi/ToolDatabase?cat=DeviceCentric&subcat=Github

17 Not all research discussed in this thesis examines natively digital objects. In those cases (such as looking at networks as storytelling devices) I resort to other methods such as multimodal analysis. I will discuss multimodal analysis in more detail in the next section.

work.

While socio-material approaches in journalism studies have developed a sensibility towards studying how digital objects matter in news work, device-centric perspectives from digital social and media research further specify what a news devices research approach would look like, notably by empiricising[18] not just the object of study but also the question of method, as I will discuss below.

First, in the news device approach, the question of possible sites of study of interactions between digital objects and news work would be steered not towards offline sites where their interactions can be observed but towards the digital devices themselves. To take an example from my own work, that would be taking GitHub instead of the newsroom as a site for observing journalism coding. Secondly, the question of methods is steered towards analytical modes afforded by the digital devices implicated in organising journalistic knowledge, experience and relations. Thirdly, and perhaps most importantly, device-oriented perspectives also shape the *scope and perspective* of study. News device perspectives explore relations and practices by attending carefully to how these are inscribed, supported and enacted by digital objects. To deploy the literary notion of "focalisation", the perspective from which a narrative is presented (Bal & Lewin, 1983; Genette, 1983), device perspectives explore how relations and situations are organised and made intelligible from the perspective of digital devices. The interactions between news work and digital objects might be empirically studied by eliciting multiple accounts and diverse perspectives of what newsroom workers themselves see as meaningful interactions and what they invoke to account for these objects and interactions. A news devices approach would explore such questions by starting with the devices themselves. Here treating news empirically implies investigating how news and journalism are handled and organised by various online devices (see also

18 While empirical social research is typically driven by the question of what the actors themselves invoke to account for society or what they treat as the social in a particular situation (Marres et al., 2018; Lynch & Woolgar, 1990), in the context of device-centred research the question of method is also displaced onto the object of study, i.e. digital devices, through the move of repurposing analytical modes inscribed in the object of study.

Rogers, 2013). Following the approaches discussed above, a news device perspective would draw attention not just to how digital devices are used for journalism but would also ask how digital devices treat, process or enact various aspects of news and journalism work, how they configure the relations between news and other social domains, and what news becomes in the context of digital devices.

A news devices approach would also develop a sensitivity and reflexivity towards how these devices participate not just in news work but also in journalism research work and methods. As Marres (2012b) argues, digital research "becomes *noticeably* a distributed accomplishment [towards which] online platforms, users, devices and informational practices actively contribute" (p. 139).

The news device approach may serve as a possible direction for the methodological questions raised in relation to journalism socio-materiality approaches discussed in section 2.1. It may also gesture towards a different path for working with digital devices as resources for the study of journalism, as an alternative to more prominent computationally-inflected and/or big data research approaches.

Finally, while the research orientation sketched out in the news device perspective would favour particular sites, modes of analysis and questions, it is intended to complement (and not replace) already existing research approaches to and methods for the study of digital journalism. I aim to illustrate what perspectives are opened up by making slight modifications to how digital devices are studied in journalism, by exploring how one might address questions about sites, modes of analysis and research questions in slightly different ways, and by developing materially-sensitive ways of thinking about method.

So far in this chapter I have introduced the conceptual, methodological and disciplinary commitments that underpin the empirical work documented in this thesis and which, combined, suggest what a news device perspective to the

study of digital journalism could entail. Given that journalism is constituted through interplays of multiple devices, the question of which devices to study and how to demarcate or delineate them as objects of study becomes important. The emphasis can be variously put on different components of the devices – such as their material affordances, resulting subjectivities, ordering mechanisms, and so on – resulting in different research contributions. These considerations will be further expanded in the next and final section of this chapter, where I describe the construction of my empirical cases.

2.3 Approaching News Devices Through Case Studies

In this section I discuss some final aspects of my research design before moving onto the empirical chapters. More specifically I introduce my cases, how they were selected and developed, and how they enable me to problematise and explore the role of digital devices in various aspects of news work from a news device perspective.

The proposal for a news device approach is developed through a series of three empirical case studies examining: (1) how networks operate as narrative devices in news and what kinds of stories they tell about collective phenomena, (2) how GitHub is used as a news device but also how it configures news work as a platform asset, and (3) how web tracking infrastructures of professional and junk news sites operate as audience commodification devices. These are intended to problematise and advance understandings of the role of these devices in journalism and its study in particular ways, namely as methods for assembling news work and research in various ways.

But before going into more detail into how these cases were selected and developed, I will introduce the case study approach which I have drawn upon in this dissertation.

2.3.1 The Case Study Approach

My empirical research is organised around case studies. While case studies are used in multiple fields of research in different configurations and for different purposes, in this section I will not review the entire literature on case studies. Rather, following Beaulieu, Scharnhorst, and Wouters' understanding of case studies as "discipline-specific ways of valuing and disregarding cases" (2007, p. 674), I will focus on the approach used in my work.

Simply put, a case study is "an in-depth study of a single whole", i.e. of a whole phenomenon of the empirical world, which may be composed of multiple elements (Morgan, 2012, p. 668). The researcher's engagement with the researched topic is generally deep and nuanced and may be realised through a combination of research methods and diverse materials. This approach is also characterised by a relative open-endedness or indeterminacy of the research angle at the beginning of the study, which is to be further specified during the research process. A case study approach typically results in a thick and complex description of the studied phenomenon (Morgan, 2012).

Following Ragin (1992), I understand cases as the outcome of a process of "casing". This concept draws attention to the fact that cases do not exist "in the wild" waiting to be discovered but they are the outcome of multiple research operations by means of which concepts and empirical materials are brought in relation and become mutually constitutive, in the sense that concepts are used to reduce the complexity of the empirical world at the same time as being constantly re-articulated through empirical evidence (Ragin, 1992). The reduction of the complexity of empirical materials should not be understood as a blanket simplification but is done in order to enrich selected aspects and support selected lines of inquiry.

In the case of my research, such casing operations might include narrower ones that have to do with decisions about the boundaries of the devices I am studying and what empirical materials to collect and analyse for each case. It may also include broader theoretical casing operations, e.g., the framing of

interactions between news work and digital objects through a news device approach, that constitutes the objects of study as socio-technical arrangements and as informing the research method. The latter is an essential framing decision with important consequences in that, while it opens some lines of inquiry, it also closes others. For example, while in Chapter 5 I examine web trackers as part of the methods through which news audience marketplaces are assembled, trackers may also be constructed as an object of surveillance studies or of user privacy measures online (for an analysis of web trackers in the service of surveillance studies see, e.g., van der Velden, 2018). In the case of the GitHub repositories that make the object of Chapter 4, while I study them as objects of platformisation, they can also be constructed as cases that enable the study of collaborative software development (see, e.g., Biazzini & Baudry, 2014).

A good way to describe how the case studies developed in this thesis make empirically-informed conceptual contributions, is through the notion of problematisation. Following Beaulieu et al. (2007), in my thesis the developed cases can be described as constituting different ways of problematising various aspects of the interactions between news and the digital through different articulations of research objects, methods of investigation and empirical materials. Beaulieu et al. (2007) describe the dominant problematisation that case studies in STS develop, as enabling research accounts to express and highlight the diversity, specificity and variation of studied technologies and knowledge making practices. This problematisation emerges in response to deterministic accounts of technological development and universalising accounts of the making of scientific knowledge. This problematisation is an important part of my inquiry into how digital devices participate in news work.

Hence, I use the case study approach not for theory testing but in order to account for phenomena of the empirical world and to develop, advance, redefine or problematise concepts informed by empirical research:

> Social scientists, at various times and in various fields, have argued that case studies are not primarily vehicles for theory testing, where this is usually taken to mean testing hypothesized relationships between

> variables. And this is not because case studies are approached theory free. Rather, so it is claimed, case studies are research in the context and service of discovery, not justification: they are for the formation of evidence-based concepts, for the development of measurement structures, the places where types are defined and kinds isolated, where phenomena might be revealed and theory developed. (Morgan, 2012, p. 671)

Such work may include identifying elements of the phenomenon that are of research interest, describing key features and aspects of the phenomenon and conceptualising its functions in the social world (Morgan, 2012).

2.3.2 Selecting Digital Objects and Developing Case Studies

A number of considerations guided the selection of objects and the construction of case studies. In addition to the conceptual and methodological outlooks discussed in the previous sections of this chapter, there are also considerations related to aspects of journalistic practices examined, considerations pertaining to digital objects and case specific methodological operations. To help guide the reader through this discussion I replicate Table 1 from the Introduction below.

Area of news work	Making narrative and storytelling	Making infrastructure and coding	Making audience
Digital object	The network graph	The coding platform	The web tracker
News device	The network as storytelling device	GitHub as connective coding device	The tracker as audience marketplace device

Table 1 (duplicate of table in introduction): Types of news work, digital objects and news devices examined in this dissertation.

From the point of view of aspects of news work covered, the thesis covers: narrative making through storytelling, infrastructure making through coding, and audience making through commodification (see Table 1). The thesis starts

in the terrain of journalistic representations, namely news and investigative stories. News stories are the most familiar, visible and accessible aspect of news work, intended to reach a wide public. Storytelling is also an aspect of news work that is regularly discussed by journalism researchers and journalists alike. Starting the empirical work with this aspect of journalism practice I think is a good way to introduce readers of this dissertation to the proposed approach, at work on a topic that is familiar and of wide interest. However, in the spirit of Becker's (1982) proposal for a sociology of art that expands the analytical frame beyond the artists and the familiar art works themselves to the wider networks of relations and practices though which they are produced and circulated, the next two empirical cases studies move towards less visible aspects of news work, namely coding and audience making. These two cases are also a response to calls that emerge in the journalism socio-materiality literature to give more sustained attention to particular aspects and areas of journalism practice.

As indicated in in section 2.1, these authors argue for approaches that emphasise things other than what they call "human-centric considerations" (Lewis & Westlund, 2015), and that explore the spatiality of news work by taking locales other than the newsroom as a site of study. For these reasons, when approaching coding, the second aspect of news work examined in this thesis, I have not focused on journalism-coders and the impacts of their work on the editorial side of news work, but on the code work itself, as available in public code repositories on the social coding platform GitHub. While the focus on code repositories of news organisations on GitHub might be seen as a reassertion of the newsroom as the main locale for news work, as the chapter will illustrate, changing the site of study provides an opportunity to re-situate the newsroom as a locale for news work. As journalism code work relies on distributions of code and contributions of participants from myriad places, it may be seen as a decentralisation of journalistic production away from the newsroom. At the same time, as the chapter points out, collective journalism code work is being recentralised outside the newsroom, on a code-sharing platform, thus complicating the critique of newsroom-centricity by adding the equally critical dimension of platform-centricity of news work.

The third empirical case investigates a less visible aspect of news work, namely the audience product. While research on the role of technology has focused mainly on the editorial sides of news work (Lewis & Westlund, 2015), in this third empirical case I focus on the role of digital objects in the business side of news work, and in assembling relations between the news industry and the digital marketing and advertising industries. The economics of news work and its relationship with the news and advertising industries have been indicated as one of the blind spots of digital journalism research (Pickard, 2017). By treating journalism as a form of digital cultural production and treating its audience marketplaces alongside those of other forms of digital content production, junk news, I am to explore what is specific to audience marketplaces of the news industry compared to other forms of digital content production. According to Boczkowski and Mitchelstein (2017), "which empirical trends are unique to online news and which ones might be shared across other domains of digital culture" (p. 17), is an important question to address. Moreover, the "fake news" scandal can be taken as an occasion to understand news as a socio-material practice that is articulated not only through the interplay between news and digital technologies as I have emphasised so far, but also in relation to other digital content producers.

From the point of view of socio-material devices, I examine three objects that participate in news work: network graphs, GitHub, and third-party web trackers. The empirical cases test digital news devices not only for their capacities to shape news work but also for their implications for the relations between different domains that news configures. Here I draw on Gillespie et al.'s (2014a) understanding of news media "not merely as messages that affect minds, but as social relations by other means" (p. 2). I also draw on Marres (2017a) who argues that it is important to trace not just the implications of digital devices for particular fields such as that of news and journalism, but to also trace how digital devices reconfigure or co-articulate relations between different social "domains", such as journalism, advertising industries, policy, computer science, data intermediary industries, and so on. Indeed, in the chapters that follow I explore how digital devices configure not just news

processes but also relations between news and other areas, from digital visual culture, to commercial online platforms, other digital content producers and the online advertising and marketing industries.

My approach to these objects and my construction of cases is also guided by calls from journalism materiality researchers to make journalism research relevant to other disciplines (Boczkowski & Mitchelstein, 2017). I do this by configuring my cases to address both implications for news and journalism and its relations with other domains, as well as to understand the digital devices themselves and how they might affect other practices.

I focus on network diagrams because over the past couple of decades, networks have become an increasingly popular way to represent all kinds of collective phenomena. But while the analytical properties of networks have received much attention (see, e.g., the area of graph theory), the narrative affordance of networks are just beginning to be explored. Studying what network diagrams contribute to journalistic stories draws attention to the narrative affordances of networks, particularly given that journalism can be taken as exemplary form of storytelling.

The second digital device examined is the code platform GitHub and how it treats public journalism code repositories. I focus on GitHub because while it is the largest online code hosting service and has received a lot of attention particularly from researchers in computer and information science and software engineering, it has been less studied in digital sociology and critical platform and software studies. By taking a device approach to the GitHub platform I aim to contribute to platform studies by developing a device-approach to understanding GitHub's platform-making processes.

The final digital object is third-party web tracking infrastructures of news industry websites and junk news sites. I focus on web tracking because this is an essential but often overlooked aspect of journalistic audience making online. Web tracking is one of the core mechanisms through which the economic model of digital businesses is being materialised. Web trackers are snippets of

third-party code embedded in news websites through which data flows are being established between media organisations, digital platforms, advertisers and other third parties, and through which the monetisation and platformisation of news being materialised. In doing so, I aim to contribute not only to media audience studies by connecting the economic underpinnings of news with its digital infrastructures, but also to infrastructural approaches to new media studies by developing the concept of trackers as audience marketplace devices in the context of news websites.

Another way to distinguish between these cases is in terms of the aspects of the device perspective, the conceptions of the digital that they articulate and the methods used. The first object, network diagrams, is a digitised one, in the sense that it is not native to the digital but rather it is a pre-digital material object that has migrated to digital media (for more on the distinction between natively digital and digitised objects, see Rogers, 2014). While network diagrams are prominently associated with online visual culture and online phenomena such as the web or social networks, these are underpinned by earlier social scientific concepts and methods such as those of social network analysis (Marres, 2017a). GitHub and web trackers may be considered natively or born digital media and objects, in that they are specific to the medium (Rogers, 2014), although these too are underpinned by established social scientific traditions, as Marres (2017a) cautions. Another consideration that informs my selection of web trackers as the final digital object is Langlois, McKelvey, et al.'s (2009) call for "a move beyond, and below the user interface" (p. 8), in the sense of approaching the digital not just through the visual aesthetics and the representations made available through the user interface but also through less visible digital objects.

Finally, the methods I use follow from the way in which the interactions between the three digital objects and the three aspects of news work are problematised through conceptual approaches, research questions and research angles.

In the case of the interactions between network diagrams and journalism

storytelling, the problematisation revolves around the narrative potential or affordances of network graphs in their original contexts of publication, journalistic articles or multimedia projects. To address the question of how these objects shape journalistic stories about collective phenomena, I use multimodal analysis on a collection of journalism pieces of different genres. Multimodal analysis is an established approach to the study of communicative texts that makes use of multiple modalities, such as graphics and written language. More formally, it refers to analytical approaches "that understand communication and representation to be more than about language, and which attend to the full range of communicational forms people use – image, gesture, gaze, posture and so on – and the relationships between them" (Jewitt, 2009a, p. 14). Given the particular focus of this chapter on networks diagrams, I also draw on insights from graph semiotics as developed by Bertin (1983), to understand how narrative readings are cued by visual properties of graphs and construct meaning from them. This method is discussed in more detail in Chapter 3. This case is device-driven in the sense that the focus is on the narrative readings cued by network graphs. Given that it does not deal with a natively digital object, the case is constructed around illustrating the device perspective with focus on the device as an object of study, and less on the device as part of method.

In the second and third cases, the devices are not just an object of study but also a methodological entry point, in that the analytical capacities inscribed in the way in which they network content, are taken as a way of knowing this content by using medium-specific digital methods (Rogers, 2013). In other words, digital devices and the traces they generate and make available, are an important source of empirical material for my analyses. In the case of the interactions between journalism coding and GitHub, the problematisation revolves around how the platform shapes journalism code and how the participation of news initiatives in the platform may be described. These problematisations are approached with a combination of methods. These include methods in software and platform studies such as interface analysis, the analysis of platform documentation such as help pages and the platform development blog, as well as tech press documenting the platform

development. They also include repurposing the platform's own methods for curating platform activities and data. This includes repository creation and update date stamps, popularity metrics and collaboration metrics. The data is collected by calling a particular end point of the GitHub API through a set of scripts developed with the Digital Methods Initiative for the purpose of this project. API calling is an increasingly common technique in digital social and media research to automate the retrieval of data from social platform APIs (Marres & Weltevrede, 2013). The studied corpus is demarcated through an expert list of news initiatives that hold a GitHub account. The method is described in detail in Chapter 4.

In the case of the interactions between third-party web trackers and mainstream and junk news sites, the problematisation revolves around the configuration of third-party web tracking infrastructures as audience marketplace devices. The aim is to explore how audience commodification can be approached from a digital device perspective. The method used in this chapter consists in repurposing the detection and classification capacities of the popular privacy protection browser extension Ghostery to identify the tracking signatures present in the source code of a collection of web pages from mainstream and junk news sites. This is done with the help of the Tracker Tracker tool developed by the Digital Methods Initiative. This can be seen as a form of cross-platform analysis that focuses on digital objects embedded in websites that create connections with third parties (Helmond, 2017; Rogers, 2018). Visual network exploration with the tool Gephi is then used to examine the presence of trackers in this set of websites and the tracking networks in which websites are embedded. Visual network exploration is a visual analytical technique that translates network structures into visual properties of network graphs to facilitate their visual exploration and interpretation (Venturini, Jacomy, & Pereira, 2015).

As discussed above, these case studies were configured to foreground and make available for analysis particular aspects: how networks shape narratives, how GitHub treats journalism code, and how audience marketplaces are configured and understood through the tracking infrastructures of websites. I

constructed these different case studies in order to attend to the diversity and specificity of the different ways in which the digital counts or makes a difference in news work: as visual culture and visualisation objects, as online platforms and as web infrastructure, as well as to illustrate different methods for studying them. By taking an orientation towards materiality, diversity and specificity of interactions between digital objects and news work, I aim to provide more nuanced accounts of how the digital participates in news and journalism. By illustrating a number of methodological approaches, I aim to develop means for such investigations to be conducted in the future.

3. Narrating Networks: Network Diagrams as Storytelling Devices in Journalism

The empirical investigation of news devices undertaken in this dissertation starts in the familiar terrain of journalistic stories.[19] The question around which this chapter is organised pertains to how we can account for narrative making or storytelling from a news device perspective.

Digital news stories, these media products which we come across daily, materialise collective phenomena, and are experienced as multimodal constructions which combine written texts, images, audio, video and other resources according to journalistic genre conventions. In the context of these stories, studies have pointed to the role of digital objects such as graphics and visualisations as important techniques that news makers mobilise to support the narration of their stories (Segel & Heer, 2010), facilitate their comprehension (Cairo, 2007; Schroeder, 2004), and attract reader attention (Utt & Pasternak, 2000). According to Latour (1990) images and graphics are powerful because they can at once stabilise phenomena and allow them to be transported across space, time and context, reasons for which he describes them as "immutable mobiles".

In this chapter I contribute to understanding the role of visualisations in knowledge construction by exploring network graphs, hereafter called networks, in the context of journalistic storytelling. I focus on this type of digital object because over the past couple of decades, network graphs, understood as sets of nodes (or vertices) connected by edges (or links), have become an increasingly popular way to represent all kinds of collective

[19] This chapter is based on an ongoing collaboration with Tommaso Venturini (French National Center for Scientific Research), Jonathan Gray (King's College London) and Mathieu Jacomy (Aalborg University) to examine network practices in the context of digital journalism, digital sociology and activism. An invitation to visit the Tow Center for Digital Journalism at Columbia University provided the initial impetus for this research. An article based on this study, of which I am the first author, was published in *Digital Journalism*, (Bounegru, Venturini, Gray, & Jacomy, 2017). Its development benefitted from input from participants in writing seminars led by Bruno Latour at the medialab at Sciences Po in Paris in December 2015 and March 2016. I am especially indebted to Paul Girard for his suggested clarifications the terminology used to describe the network story types. It also benefitted from feedback from participants in the research seminar of the University of Groningen's Centre for Media and Journalism Studies in June 2015. This chapter provides an account of this research with a focus on the news device perspective.

phenomena. They have become a core aspect of visual culture in the digital age, illustrating and attesting to the complex webs of association around us. The metaphors and imagery of networks have exploded across many fields and their analytical capacities are said to have the potential to revolutionise everything from medicine to markets to military intelligence. Advances in the computation and the visualisation of networks, the ubiquity of the internet as a network of computer networks and of the web as a network of hypertexts, have all contributed to this renaissance of networks (Barabási, 2002; Rieder, 2012; Watts, 2004).

Networks have also captured journalistic imagination. As journalists are increasingly organising their practices around the collection, analysis and narrativisation of structured and digital data (see, e.g., Anderson, 2018; Gray, Bounegru, & Chambers, 2012), the analytical and communicative potential of networks has received growing attention amongst journalism practitioners. Cases such as Josh On's map of interlocking directorates of the most powerful US companies,[20] Valdis Krebs' mapping of the terrorist network around the September 11 attacks (2002), Little Sis' and Muckety's maps of powerful people and organisations in the US,[21] have opened up imagination about the potential of network analysis and visualisation to make collective phenomena knowable journalistically.

While the mathematical properties of networks have received extensive attention for a long time (see, e.g., the research areas of graph theory and sociometry), in this chapter I follow the call of Espeland (2015) to scrutinise not just the analytical operations that make up quantification practices but also the circulation, interpretation, impact and narratives that numbers generate. Hence in this chapter I explore the stories about collective phenomena that network graph properties evoke in a series of journalistic projects and articles. Following the news device perspective described in Chapter 2, I take as an entry point to the study of journalistic storytelling not a journalism genre or media type but a defining object of digital culture, and examine its affordances

20 http://www.theyrule.net/
21 http://littlesis.org/, http://www.muckety.com/

as a storytelling device.

I start by revisiting the link between journalism and storytelling, as well as how the narrative role of networks has been discussed so far. Next I introduce my corpus and research method and proceed to describe the five narrative readings of networks I identified in a series of journalism examples. Finally, to support further research in this area I discuss methodological issues that I encountered and suggest directions for future study that I hope will advance and broaden research around this defining object of visual culture after the digital turn.

3.1 Journalism and Storytelling

Journalism has long been associated with practices of storytelling: the craft of rendering complex phenomena into narrative form (for an overview see, e.g., Kormelink & Costera Meijer, 2015; Zelizer, 2004). Schudson (1982) argues that the power of media stems not from the delivery of "facts", but rather from the development of narratives and narrative forms, i.e. conventions by means of which sequences of actions or events are being reconstructed, ordered and presented. Bell (2005) writes that "journalists are professional storytellers of our age" (p. 397).

Journalism scholars have developed a rich tradition of studying the role of narratives in journalism. One strand of studies focuses on narrative styles, forms and conventions by means of which journalists decide what becomes news and how it is structured (see, e.g., Bell, 2007; Darnton, 1975; Tuchman, 1976). A famous such style is the "inverted pyramid" of narration which we often encounter in the news genre (Bell, 2007). Recognisable through its principle of organising and ordering facts from the most to the least important, the inverted pyramid organises narration around values such as recency, novelty, human interest, conflict and focus on personalities (Bell, 2007; Dardenne, 2005).

Other studies focus on the myth-making capacities of journalism, i.e. its

capacity to create representations of the world that produce or reproduce ideological or value systems and are generative of character types such as the hero and the villain (see, e.g., Ettema & Glasser, 1988; Knight & Dean, 1982). For example, according to Ettema & Glasser (1998), in the US tradition of investigative journalism, narration often acts as an instrument for invoking morality in the service of judging civic vice, by portraying innocent citizens as victims of structural problems of the state.

In recent years the digitisation of journalism has been associated with renewed possibilities for narration, emphasising possibilities for more immersive, interactive, multimodal and participatory forms of storytelling (Pavlik, 2000, 2001). While some have cautioned about this optimistic interpretation of digital storytelling as journalism's panacea (for a discussion see, e.g., Kormelink & Meijer, 2015), empirical investigation of storytelling with digital media is ongoing. For example, augmented reality (AR) as a storytelling medium has been understood to have the potential to increase audience engagement with news (Pavlik & Bridges, 2013). Twitter has been understood as a medium for participatory storytelling around news where citizens and journalists collaboratively construct stories around news events (Papacharissi & Oliveira, 2012).

3.2 When Networks Meet Narratives

While the use of networks for data exploration and analysis has been studied extensively (see, e.g., Adamic & Adar, 2003; Andris et al., 2015), the narrative or storytelling potential of these digital objects is just beginning to receive more sustained attention from researchers.

Networks and narratives have recently come together in a number of different areas of research. One such area is that of information and communication technology and organisation studies, where narrative networks represent methodological devices for representing patterns and routines that emerge around the usage of information technologies (Pentland & Feldman, 2007;

Weeks, 2012). Another perhaps less expected direction of study is exemplified by the Narrative Networks programme set up by U.S. government's Defense Advanced Research Projects Agency (DARPA). The programme aims to explore the relationships between narratives, human cognition and behaviour in the context of international security. Another more familiar area of inquiry that brings together these two concepts is narrative theory. In this context network analysis methods have been applied to the study of literary texts in attempts to develop quantitative approaches to enrich the study of narrative texts (see, e.g., Bearman & Stovel, 2000; Moretti, 2011; Sudhahar, Fazio, Franzosi, & Cristianini, 2015).

My interest differs from these other approaches in that I am predominantly interested in the narratives about collective phenomena that networks and their properties may elicit in situated storytelling practices. This is an area of research that is just beginning to be explored (see, e.g., Bach et al., 2016; Suslik Spritzer et al., 2015; Venturini et al., 2017, 2018). Following the news device approach sensitivity towards the mutual shaping of technologies and practices (see Chapter 2), I am interested in how these digital objects are appropriated and used in situated storytelling practices, and, at the same time, in how their material affordances shape journalistic stories and translate aspects of the depicted collective phenomena. While the notion of affordance has different meanings and inflections, here I use it to account for both how the visual properties or features of networks invite particular narrative readings and to how these are activated in situated contexts of practice such as that of journalistic storytelling (for a discussion of the multiple ways in which this concept is used, see, e.g., Bucher & Helmond, 2018 and Nagy & Neff, 2015).

To distinguish my approach from the various strands of work developed around the concept of narrative networks, I propose the complementary notions of *network narratives* and *network stories*. While distinctions between "narrative" and "story" have been drawn in narrative theory (for an explanation see, e.g., Culler, 2001), in this chapter I use the terms "stories" and "narratives" interchangeably. The notion of network narrative or network story is intended to guide attention towards the narrative potential of these digital

objects, i.e. the ways in which narrative meaning may be constructed around network properties. By proposing this notion, I do not aim to suggest that network visualisations *are narratives*. Instead my aim is to draw attention, following Ryan (2004), to their potential to *possess narrativity*. Ryan characterises the distinction as follows:

> The property of 'being' a narrative can be predicated on any semiotic object produced with the intent of evoking a narrative script in the mind of the audience. 'Having narrativity,' on the other hand, means being able to evoke such a script. In addition to life itself, pictures, music, or dance can have narrativity without being narratives in a literal sense. (p. 9)

3.3 How the Analysis Was Conducted

I started the analysis by building a collection of journalism pieces which use network diagrams and concepts.[22] I collected 45 exemplary journalism pieces from a number of different sources: interviews with journalists working in this area, online repositories of journalism stories such as those associated with the Data Journalism Awards competition[23] and with the National Institute for Computer-Assisted Reporting (NICAR),[24] and specialty mailing lists such as NICAR-L[25] and Influence Mapping.[26] These pieces belong to various journalistic genres, from visual and investigative journalism, to special reports and interactives, to name just a few, and are published predominantly in digital formats. In these pieces, network graphics are integrated in multimedia packages or used to illustrate pieces of writing. The stories and network graphics which I collected were not always accessed in their original medium of publication and hence some of the elements of context in which they were

22 For the purposes of this chapter I use the terms "network diagram", "network visualisation" and "network graphic" interchangeably.
23 Accessible at: http://www.globaleditorsnetwork.org/programmes/data-journalism-awards/
24 Accessible at: https://www.ire.org/nicar/
25 Accessible at: https://www.ire.org/resource-center/listservs/subscribe-nicar-l/
26 Accessible at: https://groups.google.com/forum/#!forum/influencemapping

originally embedded are absent. However, since these elements of context are essential to the interpretation of the narrative, I only considered in this analysis those network graphics for which sufficient elements of context were available. Given the illustrative nature of this analysis I ended up focusing on 13 of the 45 pieces that were identified in the collection that I created.

My device-centred approach to the construal of narrative meaning draws on a well-documented area of research, multimodal analysis (Jewitt, 2009b; Kress & van Leeuwen, 2001). Multimodality refers to "approaches that understand communication and representation to be more than about language, and which attend to the full range of communicational forms people use – image, gesture, gaze, posture and so on – and the relationships between them" (Jewitt, 2009a, p. 14). The network stories that I selected are realised through the interaction of multiple modes, from static and interactive diagrams, to photographs, pictograms, and written language, the latter being present in various forms, such as story headline, lead, body, graphic caption, labels, legend and instructions. All these modes are meaningfully organised in the layout space of the journalistic piece. A mode is defined as "a socially shaped and culturally given resource for making meaning", utilised in representation and communication (Kress, 2009, p. 60).[27]

To facilitate the construal of narrative meaning from the point of view of the device, the network diagram, I used a model of analysis for the graphic representation of networks, an area which is only beginning to be addressed in multimodal research (see, e.g., Bateman, Wildfeuer, & Hiippala, 2017). To construe narrative meaning from network graphics, I draw on Bertin's (1983) semiotic model for the analysis of graphics, and on Venturini et al.'s (2015) framework for the visual analysis of networks. Bertin (1983) proposes that the reading of a graph largely consists of constructing correspondences around a central notion called the invariant, characterised by means of a series of visual variables, such as size, shape, colour and texture (p. 140-141). This reading occurs in stages, such as identifying the informational notions or concepts

[27] For a more extensive discussion of semiotic modes see, e.g., Bateman (2011) and Hiippala (2014).

expressed through the graph and detecting the visual variables through which these notions are represented, before correspondences are being established between graph components. Crucial to the creation of such correspondences and thus to the construal of narrative meaning is a *question*, conscious or not, which guides the reader's perception towards the relevant associations. Venturini et al. (2015) apply Bertin's (1983) model to network graphics and propose a framework for constructing meaning around three visual variables of network graphs: node position, node size and node hue. According to this model, visual patterns corresponding to network properties such as clustering (Watts & Strogatz, 1998) and structural holes (Burt, 1992), may be read from the spatialisation of the network, i.e. the disposition and density of nodes in the graph. The property of an actor to be an authority, a hub or a bridge (Barabási, 2002; Granovetter, 1973), i.e. their centrality in the network, can be read from the position of nodes in the network, as well as from their size. Actor typology may be read from node colour or shape. In the next section, I illustrate how I applied these models to the analysis of multimodal journalistic stories with a network graphical component.

Another essential aspect in narrative construction and interpretation in our corpus is the genre of journalistic storytelling. Following the convention of this genre that the key points are presented at the beginning of the story, particularly in the headline and the lead paragraph (Bell, 2007), where applicable, I have taken guidance in the formulation of the question that the network graphic invites the reader to ask, and hence of the story which it evokes, from the headline and lead of the journalistic piece, as well as from the graph elements addressed in textual form in the body of the story. Finally, the socio-cultural knowledge of the reader also plays an important role.

Drawing on these models, I take a number of steps in my analysis. Given my interest to explore narrative from the point of view of the digital object, the starting point of my analysis is to identify the notions, themes or concepts central to the journalism piece that are expressed through the network graphic, or what in multimodal research may be termed the diagrammatic mode (Alshwaikh, 2009). Next I detect the prominent visual attributes of the

network visualisation which guide readers' perceptions and by means of which narrative meaning is cued. The narrative view is then further specified and qualified by means of the textual elements in close proximity to the graph, most importantly its caption and the title, lead paragraph and body of the journalistic piece in which it is embedded, also mobilising more broadly our knowledge of the world and of the journalistic genre. I verbally formulate the narrative views which I construe and where applicable draw associations between them and the network concepts or properties which they deploy or with which they resonate.

To move from narrative readings to narrative reading types I use an emergent categorisation approach whereby categories of narrative readings are being construed by identifying repetitive story types or patterns in my collection of journalism pieces. Given the qualitative nature of my study focused on demonstrating the narrative potential of networks, I do not exhaustively analyse my collection of journalism pieces. Instead I focus on the construal of an illustrative set of narrative reading types, each of which is illustrated with the analysis of three examples.

3.4 Five Narrative Readings of Networks

This qualitative analysis shows that there are recurring narrative reading types that networks elicit and that these narrative readings resonate with or deploy a number of network properties or concepts. In this section I discuss five such examples of narrative readings and illustrate them with the analysis of three cases where they occur. As the analysis below shows, multiple narrative readings may be construed from a journalistic piece, which means that a journalism project or article may be discussed in relation to multiple categories. The construal of multiple narrative views occurs particularly in the case of pieces which deploy a mode which could be tentatively termed "dynamic diagram". This composite mode incorporates a strong interactive component which helps to connect different narrative sequences and enables the narrative to progress. Unlike a static diagram, the dynamic network diagram enables the

user to explore and manipulate the graph display through a variety of interactive devices and techniques. Given the illustrative nature of my proposed typology of narrative readings, I do not undertake to exhaust all of the narrative readings of a journalistic project. Instead I focus on the ones that best illustrate the narrative views under examination.

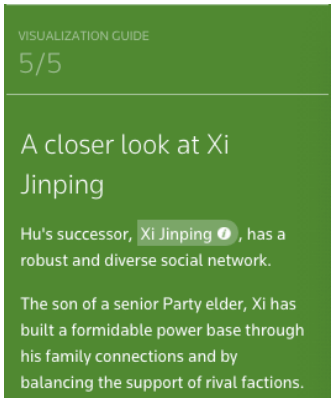
3.4.1 Exploring Associations around Single Actors

In this category I grouped narrative views that depict the network of entities around a single actor. This category may be interpreted as evoking a particular type of network graph, often called "ego-network". One typical characteristic of ego-networks is the depiction of relationships around a given social unit, referred to as the *ego*, resulting in "a mini-network or immediate neighbourhood surrounding the ego" (Freeman, 1982, p. 291).

I also distinguish this category from the second narrative view which depicts key players, in that the nodes that play the role of the ego in this category of stories are not necessarily well-connected ones (authorities or hubs), which is where the emphasis lies in the "detecting key players" category in the next section. In borderline cases, where it has been difficult to place a story in one category or another, I returned to the written language mode and particularly to the journalism piece headline and lead as well as to the graphic headline to identify additional cues for the construal of the narrative view.

I will discuss three cases of narrative views which I have construed as "exploring associations around single actors" in this collection of journalism pieces. Across all these cases, crucial to the construal of the selected narrative reading is the interactive component of the dynamic diagrammatic mode. Interactive techniques such as clicking or mouse-over network elements to reveal details-on-demand, changing the appearance of the arrow cursor into a hand over active areas, implicit instruction through visualisation guides and explicit instruction through mouse-hover over elements of the graph, guide the reader to select single nodes in order to explore their networks of associations.

Thomson Reuters' "Connected China" project (February 2013)[28] is an interactive multimedia website dedicated to "tracking thousands of people, institutions and connections that form China's elite power structure" by depicting networks of familial, political and social ties amongst members of its governing structures. Of the multiple narrative readings elicited by the dynamic diagrammatic mode, in this section I describe a prominent one, the exploration of associations around single actors. This narrative reading draws on multiple modes. In the written language and layout modes the importance of the ego is expressed through elements such as the guide to the reading of the visualisation (see Figure 1), the textual explainers describing each individual in the network and the labels identifying them by name. In the diagrammatic mode, an important cue is the prominent positioning of the ego in the graph, its representation through a photograph of the individual whom it represents, and the disposition of the related nodes in a semicircle around it (see Figure 1(b)).



[28] Available at: http://china.fathom.info/

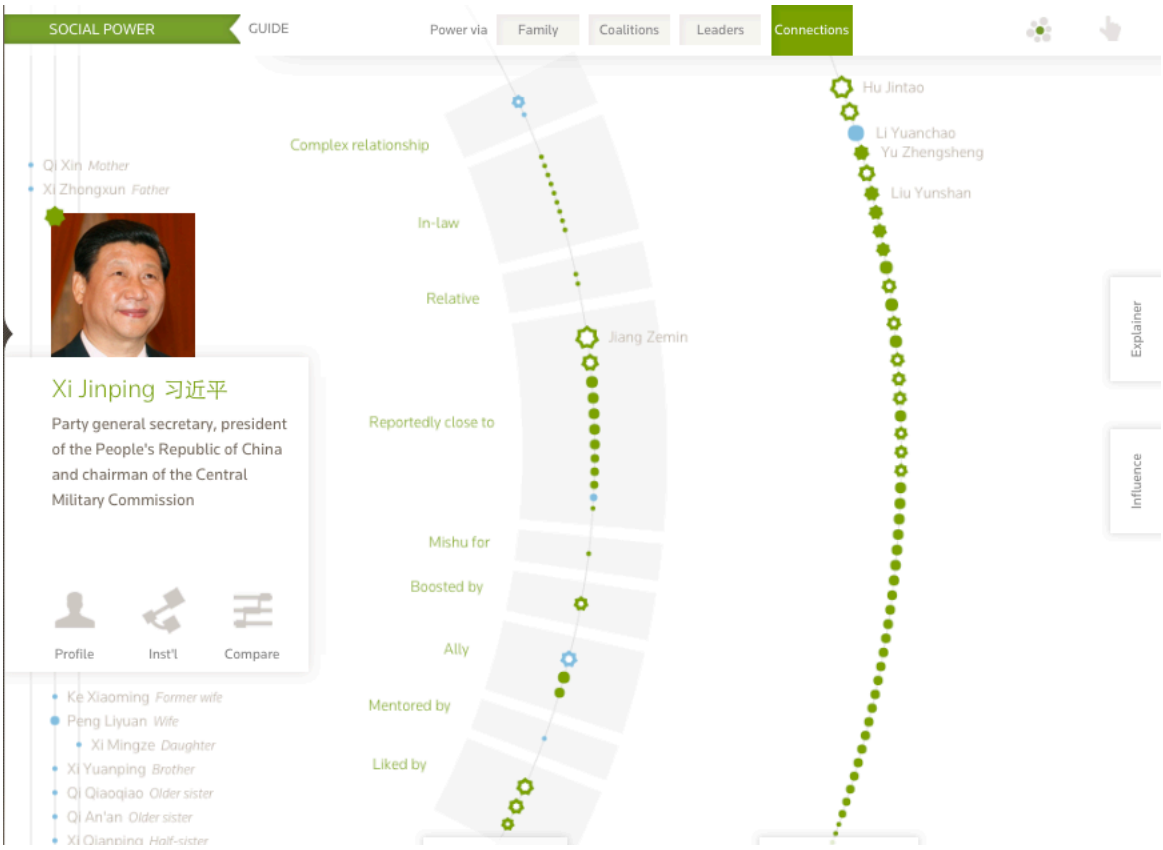


Figure 1: "Connected China", Thomson Reuters, February 2013. (a) Part five of the visualisation guide which provides instructions for navigating the "Social Power" section of the multimedia website. (b) Network visualisation depicting the network of associations around a selected government official in China.

In the interactive micro-site "WESD Web of Connections" published by the *Statesman Journal* (n.d.)[29], the Willamette Educational Service District's network of connections is a multi-layered story about corruption, oversight and abuse undertaken by a regional education service agency in the United States. Multiple modes participate in the construction of this narrative reading. In the default view of the interactive network graphic, the Willamette Educational Service District is set as the ego but users can move any other node to the centre of the graph by clicking on it. The written language mode, more specifically the graph title, instructions and labels, interacts with the dynamic diagrammatic mode and its spatial disposition of nodes and edges to guide the interpretation of the network around a single actor (Figure 2). The primary visual property through which the actor's network is realised is the spatial

29 Available at: http://community.statesmanjournal.com/news/wesd/web/

disposition of the ego at the centre of a radial graph, with edges radiating from it and connecting it to nodes representing suspect deals, administrative structures and the individuals, organisations and addresses associated with these.

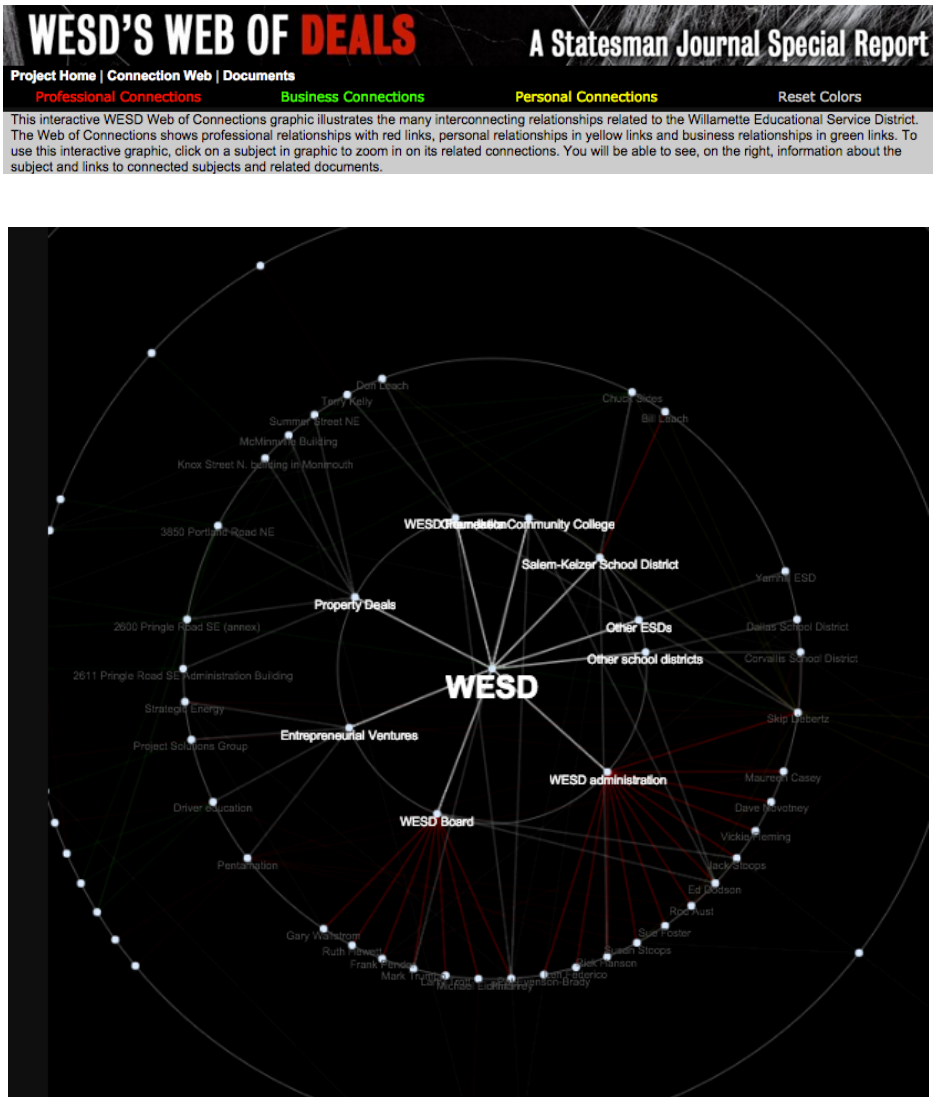
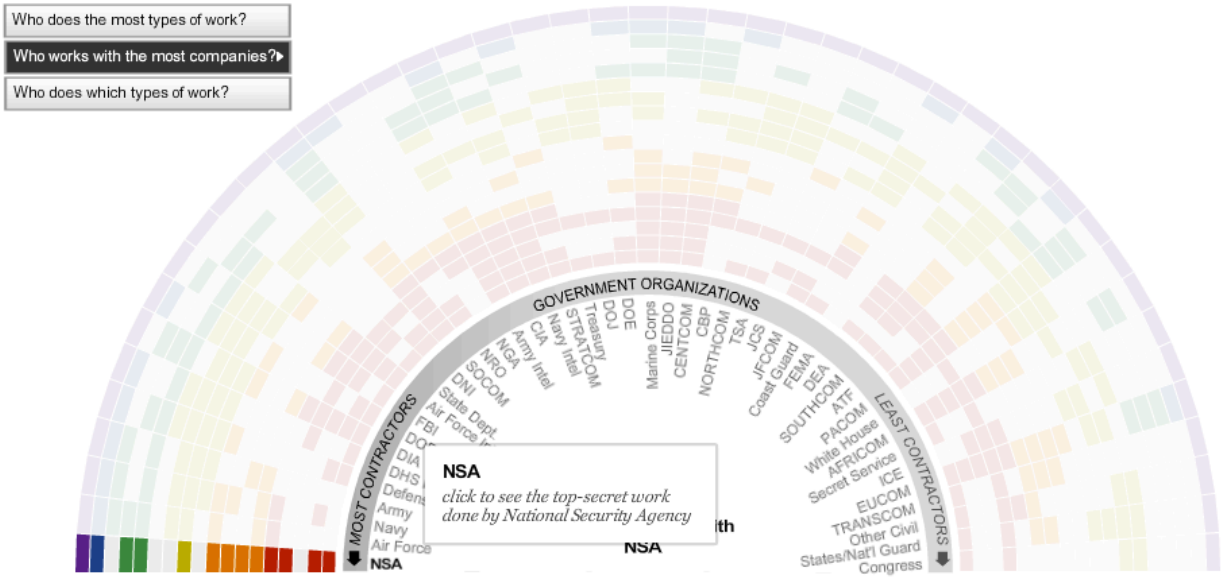


Figure 2: "WESD Web of Connections", Statesman Journal, n.d. (a) Special report title and instructions for navigating the interactive graphic. (b) Default graphic view depicting the Willamette Educational Service District and its network of connections.

Finally, Washington Post's "Top Secret America",[30] an investigative project published on July 19, 2010, enables the user to explore the network of types of work conducted by and companies contracted by each of the 45 government

[30] Available at: http://projects.washingtonpost.com/top-secret-america/

organisations that make up the national security program of the United States. This narrative reading is cued by the written language and dynamic diagrammatic modes. An interactive device, an information box appearing upon hovering over elements of the visualisation, guides the reader to click on the name of an agency in order to explore its work areas and contractors (see Figure 3(a)). The ego is cued in the diagrammatic mode by its spatial disposition at the centre of the graph, as well as the size of the node representing it. Its network of entities is depicted through a series of circles of varying sizes and blocks of varying colours. The ego and its network of elements are further qualified in the written language mode through the graphic's caption, the information box and the node label (see Figure 3(b)).



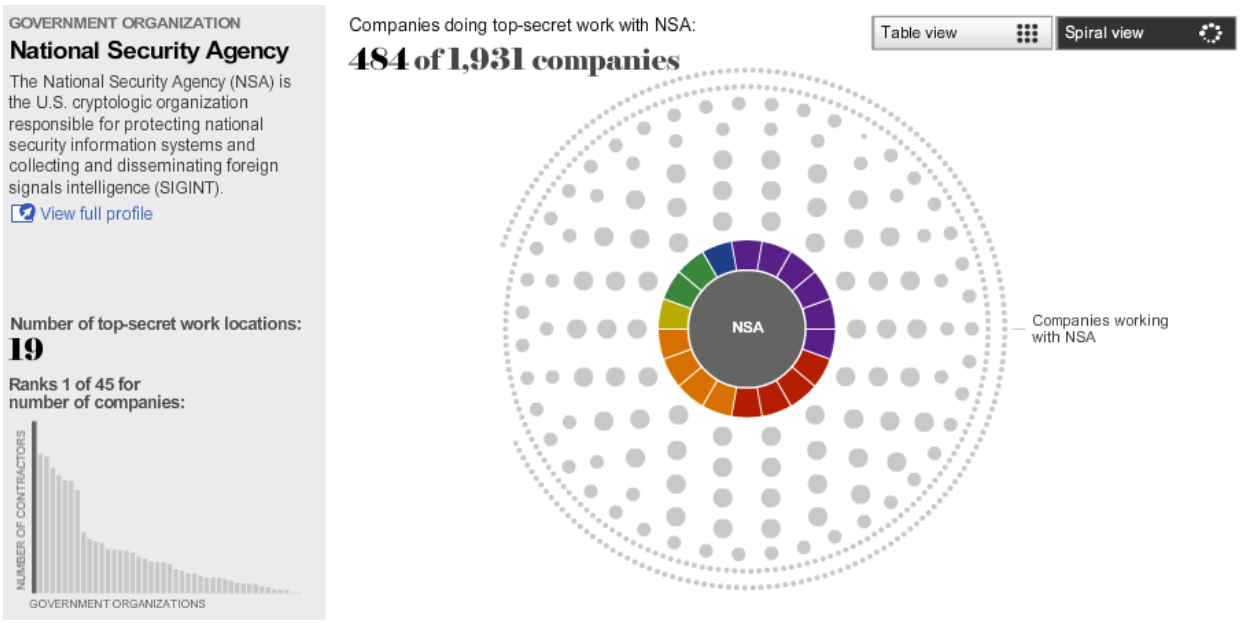


Figure 3: "Top Secret America", Washington Post, July 19, 2010. (a) Default view of the "Explore Connections" section of the investigation with information box inviting the reader to click on an entity to see its network of connections. (b) Network of contractors and areas of work around selected government agency.

3.4.2 Detecting Key Players

In this category I grouped narrative views that depicted key actors based on the number of connections with other nodes. The focus of these network stories is on the density of connections around one or several central nodes. This category deploys the network property of power law distribution (Barabási, 2002). The Pareto or power law distribution of connectivity indicates the concentration of a large majority of connections around a small minority of nodes. Such nodes are called authorities or hubs, to show that they receive or spawn an unusually large number of links. Such a distribution of associations can be observed in the topology of the web, where a few webpages receive millions of incoming links, whereas close to 90% of all webpages receive ten or less incoming links (Barabási, 2002). This is also identified as a property of many other types of self-organizing networks in biology, economy and society.

I will illustrate this narrative view with three examples. The first one is the *New*

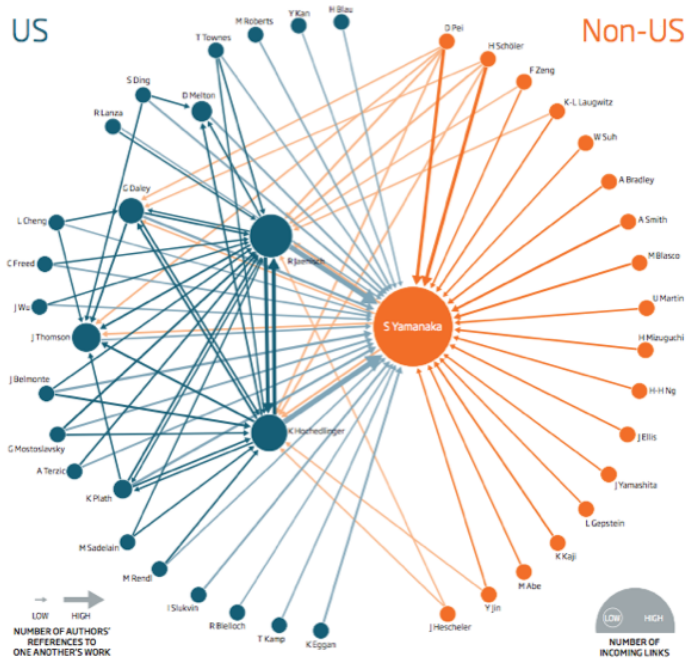
Scientist's "The Stem Cell Wars" special report[31] on citation practices in stem cell science, published on June 12, 2010. In this example the written language mode contributes to the narrative reading in multiple ways. The graphic caption guides the reading of the graphic around the detection of "influential players" (see Figure 4(a)). The information box "The strongest link" also cues this reading in its first sentence: "Shinya Yamanaka of Kyoto University in Japan is the dominant scientist in cellular reprogramming" (*New Scientist*, June 12, 2010; see Figure 4(b)). The most cited scientists, the Japanese Yamanaka and the U.S. based Jaenisch and Hochedlinger, are also discussed extensively in the body of the article. In the static network diagram, the node size as well as the depiction of edges as arrows pointing towards the most cited scientists support this narrative reading. The edges here represent citations received by papers authored by the scientists depicted as nodes in the graph.

[31] Accessible at: http://www.peteraldhous.com/Articles/The_stem_cell_wars.pdf

THE STEM CELL WARS

When a Nobel prize is up for grabs, do scientists across the globe compete on a level playing field? **Peter Aldhous** investigates

The most influential players in cellular reprogramming are revealed by recording how many times the scientists have referred to each other's work. Each link shows where one researcher cited another four or more times in papers in leading journals (for analysis, see "The strongest link", below right)



THE STRONGEST LINK

Shinya Yamanaka of Kyoto University in Japan is the dominant scientist in cellular reprogramming, but he has stiff competition from a well-linked group of US-based researchers.

To map influence in the field, *New Scientist* constructed a social network diagram (left) based on citations, the references to each scientist's work by their peers. Citations are a measure of a researcher's impact and influence, and are sometimes used to help make decisions on promotions. They can also provide a snapshot of who's who in a field.

Assisted by Henri Schildt of Imperial College London, a specialist in citation analysis, we looked at references between 148 papers published in prominent journals since 2006 – drawing links where the authors cited one another's work four or more times.

Yamanaka's research is referred to by just about everyone. But there are no such links between other scientists outside of the US, and no links to them from any of the US-based researchers. In the US, there is a richer web of connections.

In large part, this reflects the greater number of papers in our sample from scientists in the US. But another tie links the best-connected researchers in the US: the Boston area. Rudolf Jaenisch at the Massachusetts Institute of Technology is Yamanaka's strongest rival, and two of the other main players – Konrad Hochedlinger and Kathrin Plath – used to work in his lab. Plath has moved to the University of California, Los Angeles, while Hochedlinger remains nearby at the Harvard Stem Cell Institute, which also hosts the labs of George Daley and Doug Melton.

The outsider is James Thomson at the University of Wisconsin-Madison, who first isolated human embryonic stem cells in 1998. He owes his prominence in this network to winning the race, in a tie with Yamanaka, to make human iPS cells.

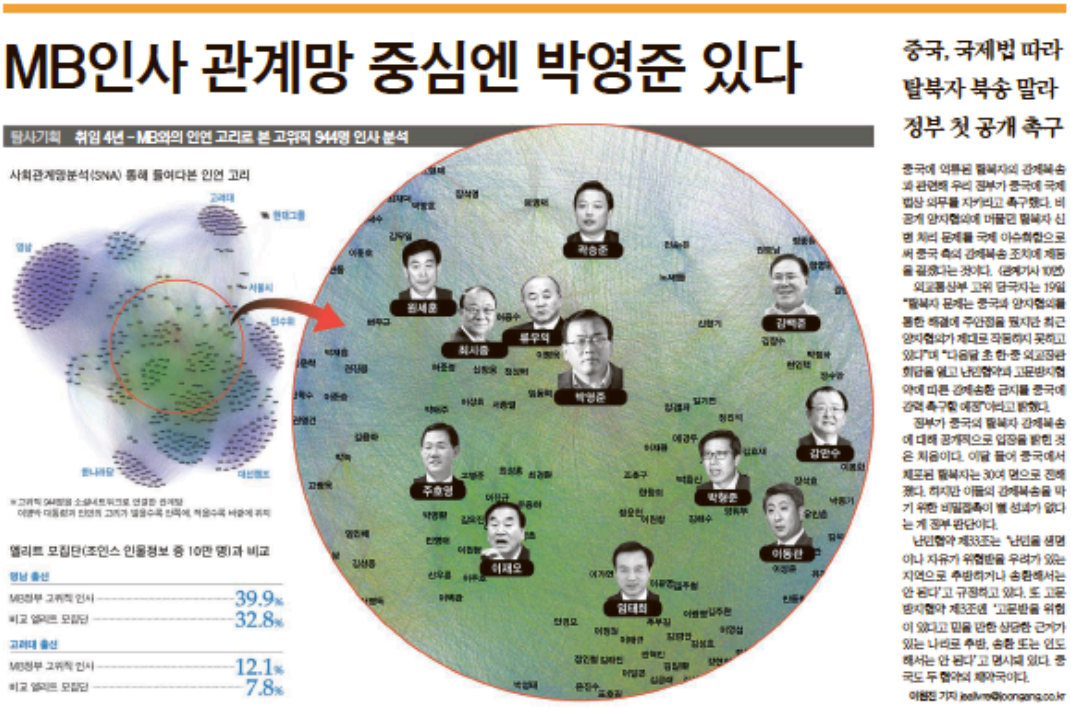
Figure 4: "The Stem Cell Wars", New Scientist. June 12, 2010. (a) Article title and subtitle, graphic caption and network visualisation of citations in stem cell science. (b) Information box describing how the

network analysis was conducted, as well as its key findings about the influential players.

The article series entitled "Park Young-Joon at the Center of President Lee Myung-bak's Human Resources Network", published on January 20, 2002, is a *JoongAng Ilbo* investigation into the social network of senior officials around South Korea's president, Lee Myung-bak. In the written language mode, the headline of the article,[32] the caption of the graphic,[33] and the body of the article all point towards the strength of associations of government official Park Young-joon to other top officials with links to the president. These links include whether they share the president's hometown, whether they graduated from the same university, whether they served on the president's campaign and a number of other business and political ties. In the static diagram the position of the node representing the most connected government official at the centre of the graph and the size of the icon representing him supports such a narrative reading (Figure 5).

[32] In translation: "In this social network among the 944 senior officials, Park Young-joon is located at the center. An official is positioned toward the center when he or she has more links to President Lee". Source: NICAR Stories Database, story no. 25691.

[33] In translation: "In this social network among the 944 senior officials, Park Young-joon is located at the center. An official is positioned toward the center when he or she has more links to President Lee". Source: NICAR Stories Database, story no. 25691.


MB인사 관계망 중심엔 박영준 있다

탐사기획 취임 4년 – MB와의 인연 고리로 본 고위직 944명 인사 분석

사회관계망분석(SNA) 통해 들여다본 인연 고리

엘리트 모집단(조인스 인물정보 중 10만 명)과 비교

39.9%
32.8%
12.1%
7.8%

중국, 국제법 따라 탈북자 북송 말라 정부 첫 공개 촉구

Figure 5: "Park Young-Joon at the Center of President Lee Myung-Bak's Human Resources Network", JoongAng Ilbo, January 20, 2002. Article title, graphic caption and network visualisation depicting government officials central to the president's network.

My final example in this section is the "Flip Investigation"[34] published by the *Sarasota Herald Tribune* in 2009, examining real-estate fraud in the state of Florida. While the dynamic diagrammatic mode invites multiple readings, in this section I focus on the "detecting key players" narrative view, which is prominently cued in the article series accompanying the interactive network visualisation. Here the key player is construed as a villain situated at the centre of a network of fraudulent transactions known as property flips. The narrative reading is signalled in the written language and photograph modes, with a full text article in the series, published on July 21, carrying the evocative title of "The King of the Sarasota Flip", being dedicated to the investigation of the key player, Craig Adams (see Figure 6(a)).[35] In the diagrammatic mode, the narrative reading is realised through the default frame of the "Network" view of the interactive graph depicting Craig Adams at the centre of the flip deals network, comprising of other real-estate professionals, "flippers", orchestrators

34 Accessible at: http://projects.heraldtribune.com/investigateflip/investigateflip.html
35 Accessible at: http://www.heraldtribune.com/article/20090721/ARTICLE/907211055

as well as victims (see Figure 6(b)). The topology of the network comes very close to the star-shaped "winner takes all" topology (Barabási, 2002), in which the central hub is connected to all the other nodes.

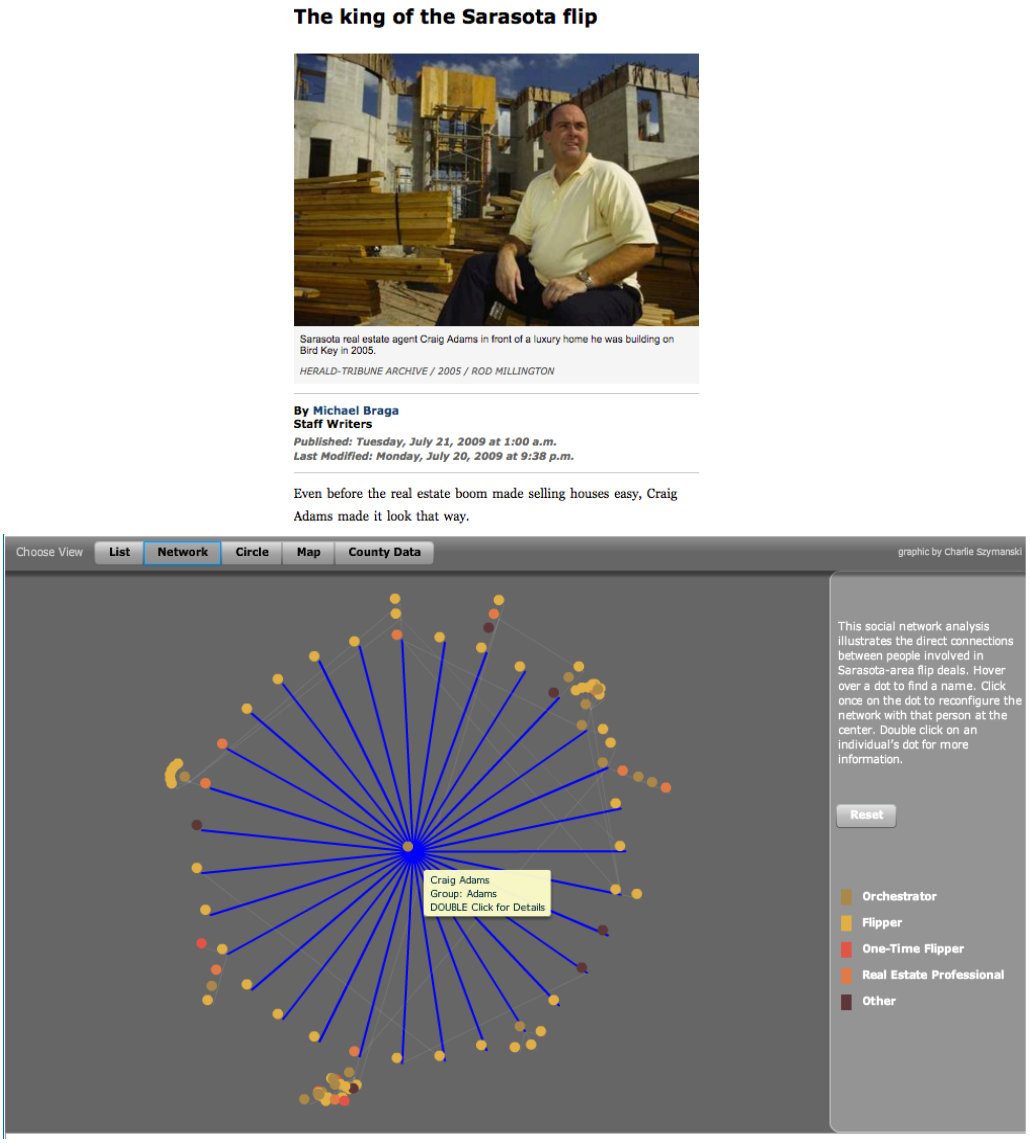


Figure 6: "The Flip Investigation", Sarasota Herald Tribune, 2009. (a) Title and picture from article in the Flip Investigation series investigating real estate agent Craig Adams. (b) Network visualisation depicting Craig Adams at the centre of a property-flipping network in Florida.

3.4.3 Mapping Alliances and Oppositions

In this category I grouped narrative views that depict associations of nodes as well as the absence of associations between groupings of nodes. This category

deploys the network property of clustering, which measures the varying density of connections between nodes (Watts & Strogatz, 1998) as well as the property of structural holes (Burt, 1992). A cluster is thus a collection of nodes that are more densely connected among each other than to the rest of the network. A cluster is visually displayed by the spatial disposition of such nodes in proximity to each other in a network graph. According to Burt (1992), the concept of structural hole refers to the "separation between nonredundant contacts" where nonredundant contacts are nodes that do not share connections (p. 18). In the visual representation of a network, the clusters may be construed as alliances between actors and the absence of connections or structural holes may be construed as opposition or lack of allegiance (Venturini et al., 2015).

In what follows I will illustrate the construal of this narrative view with three examples from the political domain. *Le Monde*'s piece "2007-2011: La Cartographie de la Blogosphère Politique",[36] is an interactive graphic published on July 4, 2011 which represents the linkages amongst the French political blogosphere between 2007 and 2011. In the diagrammatic mode, the visual property through which the reading of alliances or coalitions between actors is realised is the spatial disposition of nodes in the graph into clusters. The clusters are identified through the density of associations amongst nodes, and through their colour (see Figure 7). Alliances may also be read not only between nodes but also between clusters. Such alliances can be read from the position of clusters in relation to one another. Clusters that are closer and share more links (such as the greens and the left bloggers' clusters in the 2011 map) can be read as allies whereas the absence of links or the presence of structural holes between two clusters cues the reading of opposition (as in the case of the sparse connections between the extreme right on the one hand, and greens and left bloggers on the other in the 2011 map).

The journalistic genre convention that the most important information is presented at the beginning of the piece provides essential guidance, as the

36 In translation: "2007-2011: The Cartography of the Political Blogosphere"; accessible at: http://www.lemonde.fr/election-presidentielle-2012/visuel/2011/07/04/la-cartographie-de-la-blogosphere-politique_1544714_1471069.html

network of clusters is the default view of the graph. The alliances and oppositions are further qualified through the written language mode. The title or headline of the piece identifies the graph as a map of the French political blogosphere. An interactive component, the navigation menu on the left-hand side of the graph identifies the political factions textually through their labels ("gauche", "extreme gauche", "ecologie politique", "centre", "droite", and "extreme droite")[37] and pictorially through the colour of the icon associated with each political faction, which in turn corresponds to the colour of the clusters on the map. The reading guide which accompanies the piece provides further guidance on the interpretation of the position of the nodes in relation to each other:

> Ce type de placement permet de rendre visibles les dynamiques communautaires et le fait que certains sites échangent fortement entre eux et beaucoup moins avec les autres sites de la carte (qui seront donc plus éloignés) (Le Monde, 4 July 2011).[38]

37 In translation "left", "extreme left", "greens", "right" and "extreme right" respectively.
38 In translation: "This type of disposition helps to make visible community dynamics and the fact that some sites exchange more links among themselves than with the other sites on the map (which will therefore be more distant)"

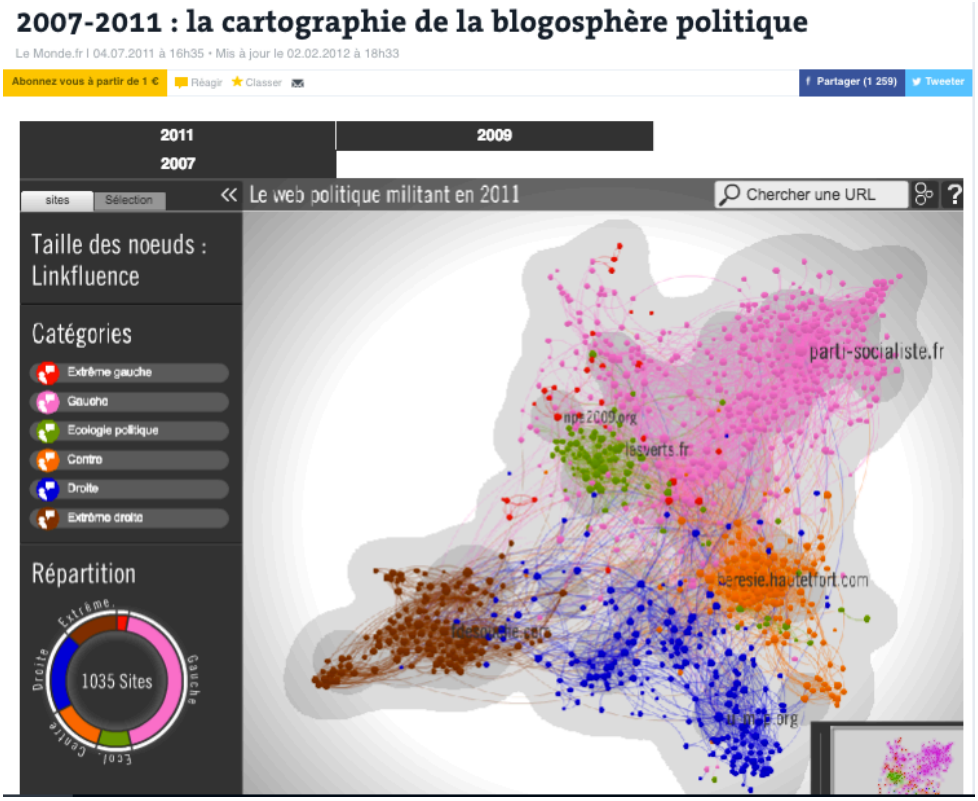


Figure 7: "2007-2011: La Cartographie de la Blogosphère Politique", Le Monde. July 4, 2011. Title of journalistic piece and network graphic view of the French political blogosphere in 2011.

Global News' piece "Visualising the Split on Toronto City Council", published on March 20, 2012,[39] tells the story of the growing divergence between Toronto council members between 2010 and 2012. As with the previous example, the alliances and oppositions are rendered in the diagrammatic mode through the spatial disposition of the nodes in separate clusters. The clusters are represented through different colours (red and blue) and are bridged by a few nodes in purple (see Figure 8). The narrative reading is further qualified as representing the alliances and oppositions around voting on the Toronto City Council through the written language mode, more specifically the headline of the article ("Visualizing the Split on Toronto City Council"), and the body of the article which identifies the factions as follows: "blue for the group around the mayor, red for the opposition, and purple for centrist or unaffiliated councillors" (*Global News*, 20 March 2012). The principle of association and

39 Archived version accessible at: https://web.archive.org/web/20120507224159/http://www.globalnews.ca/topics/torontocouncildiagrams/dec_2010/index.html

opposition is also identified in the written language mode: "Councillors who tend to vote together will be clustered closely together in the graphic" (*Global News*, 20 March 2012). The coalitions and oppositions are calculated and represented at different moments in time. This constitutes another narrative view type, which I discuss in the fourth category below.

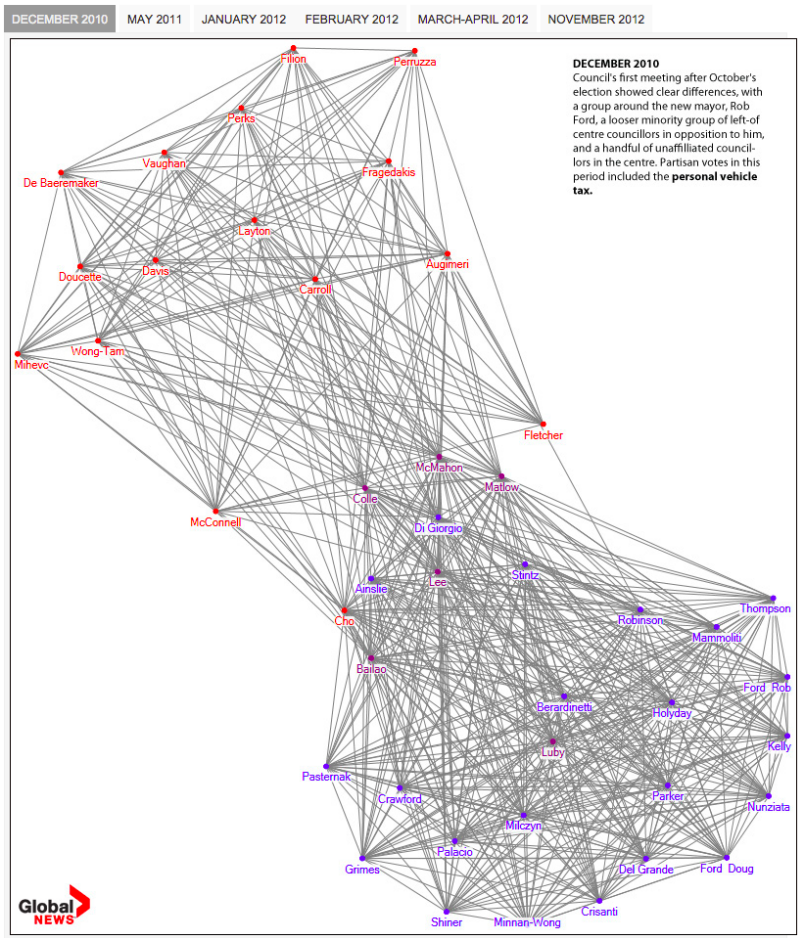


Figure 8: "Visualizing the Split on Toronto City Council. Global News, March 20, 2012. "December 2010" graphic panel depicting the group around the mayor (blue) and the opposition (red) and a few unaffiliated councillors positioned between the two clusters (purple).

Lastly, the *Le Monde* piece, "Mariage Gay: L'Opposition Soigne ses Amendements", published on January 31, 2013,[40] investigates coalitions formed around amendments to the gay marriage law proposal in France. The diagrammatic mode signals the reading of coalitions through the spatial

40 Accessible at: http://www.lemonde.fr/societe/article/2013/01/31/mariage-pour-tous-l-opposition-soigne-ses-amendements_1825467_3224.html. In translation the title reads: "Gay Marriage: The Opposition Signs its Amendments".

disposition of the nodes in several clusters in the interactive network graph which concludes the article. Political parties are represented through different colours and largely correspond to the clusters representing coalitions around the signing of legislative amendments (see Figure 9). The reading of coalitions is further specified in the headline and body of the article, as representing the alliances around the signing of amendments to the gay marriage law. The alliances are shown to follow ideological lines, with co-signing occurring largely within the boundaries of the political party, with the exception of the right-wing UMP (Union for a Popular Movement), which co-signs with independent parliament members as well as with the centrist UDI (Union of Democrats and Independents). The absence of connections between centre-left parties such as the RRDP (Radical, Republican, Democratic and Progressist) and ECOLO (Ecologists), as well as between centre-left and right-wing in the diagram can be construed as opposition. The principle of association between nodes in the graphic is also identified in the written language mode as follows:

> Chaque cercle représente un parlementaire. Lorsqu'un député a signé ou cosigné un amendement avec un autre, un lien les relie. Les cercles sont ensuite placés de telle sorte que les cercles attachés les uns aux autres se rapprochent et que ceux qui ne sont pas attachés s'éloignent. La taille du cercle varie en fonction du nombre d'amendements signés par le député. (*Le Monde*, January 31, 2013)[41]

41 In translation: "Each circle represents a member of the Parliament. When a member has signed or co-signed an amendment with another, a link between them is established. The circles are then placed so that the circles linked to each other are closer. The size of the circle is proportional to the number of amendments signed by the member of Parliament".

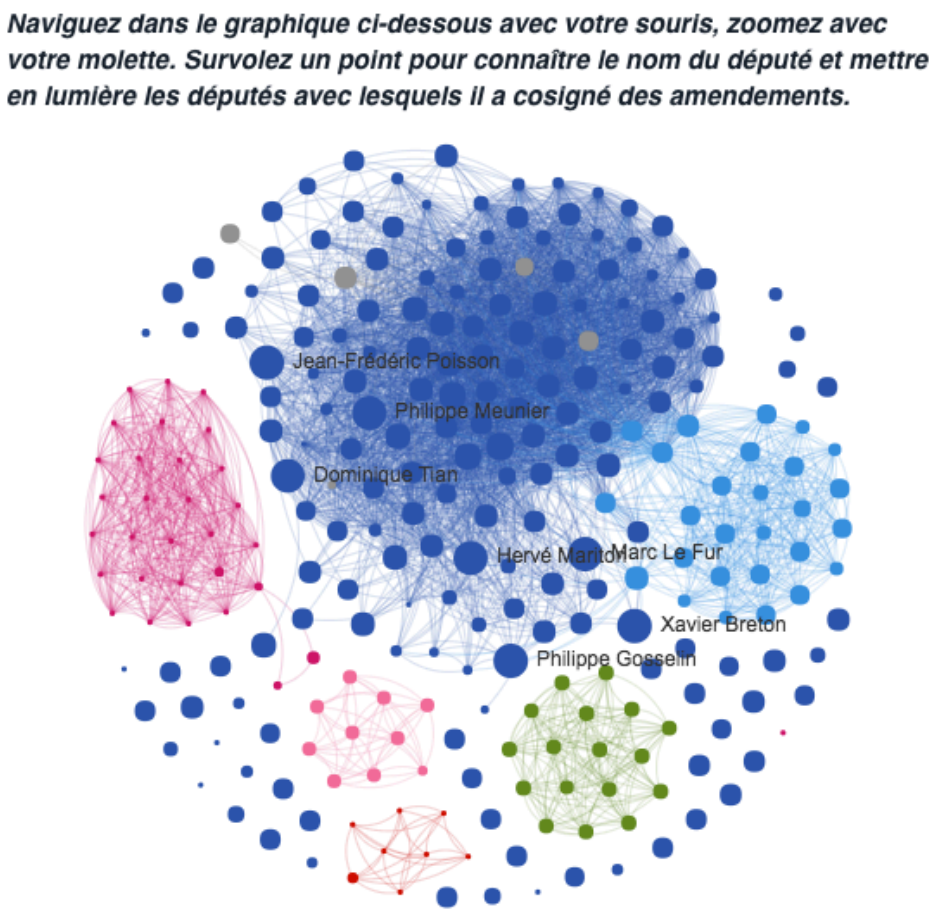


Figure 9: "Mariage Gay: L'Opposition Soigne ses Amendements", Le Monde. January 31, 2013. Instructions for navigating the graphic and network visualisation depicting coalitions around amendments to the gay marriage law proposal in France.

3.4.4 Exploring the Evolution of Associations over Time

In this category I grouped narrative views formed around a temporal dimension and which show the transformation of associations of actors over time. This category deploys a property that is common to real-world networks, namely that they are dynamic systems whose composition and topology are subject to change over time (Barabási, 2002). In what follows I will illustrate this narrative view with three examples, two of which have been encountered in previous categories as well. In the case of all three examples, essential to the construal of this narrative reading are the composite modes of the dynamic diagram or the animated diagram (as is the case of the third example) with their interactive component. Interactivity is realised through devices such as a

navigation menu indicating different moments in time in the case of the first two examples and a time bar or progress bar in the case of the final example, all of which enable the reader to navigate across visualisation panels.

In *Le Monde*'s "2007-2011: La Cartographie de la Blogosphère Politique" (July 4, 2011), I identify a second narrative view, which I call "exploring the evolution of associations over time". The title or headline of the piece identifies the time period covered by the mapping: 2007 to 2011. An interactive device, the navigation menu (containing three buttons: "2011", "2009", and "2007") and the caption of each visualisation panel ("le web politique militant en 2007", 2009 and 2011 respectively) specify the particular year represented by the panel. The narrative view is also cued pictorially by the representation of network dynamics at two-year intervals through three maps, which the user can navigate from a menu at the top of the interactive graphic (see Figure 10).

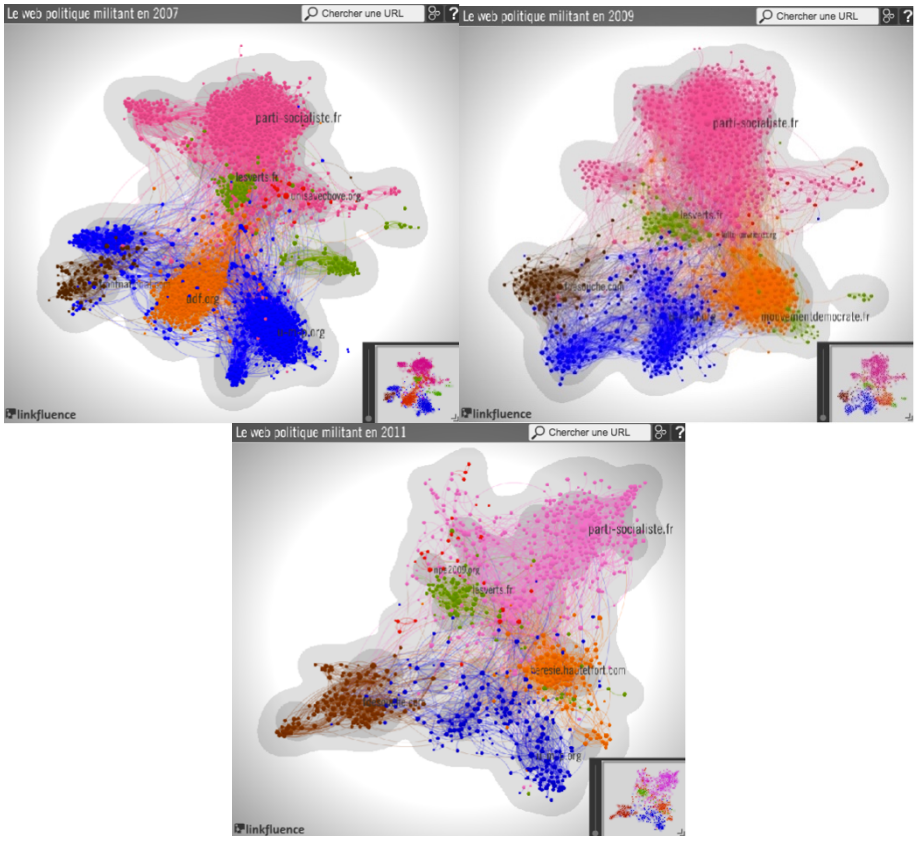


Figure 10: "2007-2011: La Cartographie de la Blogosphère Politique", Le Monde. July 4, 2011. (a), (b) and (c). Network mapping of the French political blogosphere at three moments in time: 2007, 2009 and 2011.

The narrative view "exploring the evolution of associations over time" is similarly construed in the piece "Visualising the Split on Toronto City Council" (*Global News*, March 20, 2012). One important interactive element which cues this narrative reading is the navigation menu which enables the user to switch between representations of the network at various moments in time (December 2010, May 2011, January 2012, February 2012, March-April 2012), as well as the caption of each graph, which indicates the moment in time which it depicts (see Figure 11). The longitudinal or temporal dimension is further qualified in the body of the article, where the key shifts in network dynamics at different moments in time are being analysed.

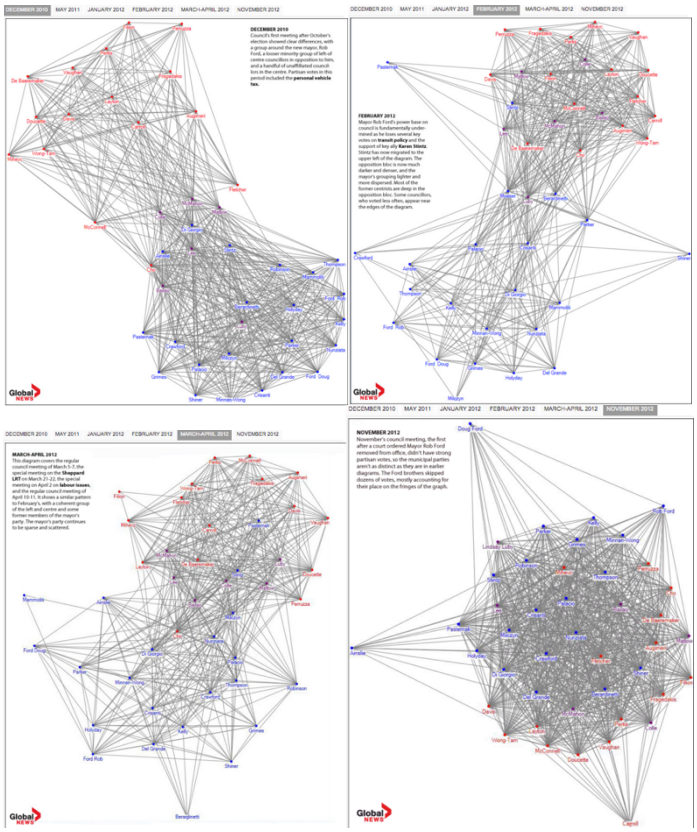


Figure 11: "Visualizing the Split on Toronto City Council. Global News, March 20, 2012. (a), (b) and (c). Network visualisation of voting coalitions on the Toronto City Council at three moments in time: December 2010, January 2012 and February 2012.

Lastly, in the *Guardian*'s interactive piece "How Riot Rumours Spread on

Twitter" (December 7, 2011),[42] an investigation into the emergence and correction of misinformation pertaining to the 2011 UK riots on Twitter, this narrative view is realised primarily with the participation of the dynamic diagrammatic mode, the written language mode and the layout mode. An interactive object, the time bar or progress bar present on the web page dedicated to each rumour, enables the reader to follow the development of the rumour from its inception until its death. The diagrammatic mode signals the similarity of tweets through their spatial proximity in the graph, as well as their position in relation to the rumour (support, opposition, questioning and commentary), through the property of colour (see Figure 12). This narrative reading is further qualified through textual annotation highlighting key events in the unfolding of the rumour as well as through the intensity of the nodes' colour, whereby lighter tones are used to represent more recent tweets and darker tones are used to represent older tweets.

[42] Accessible at: http://www.theguardian.com/uk/interactive/2011/dec/07/london-riots-twitter

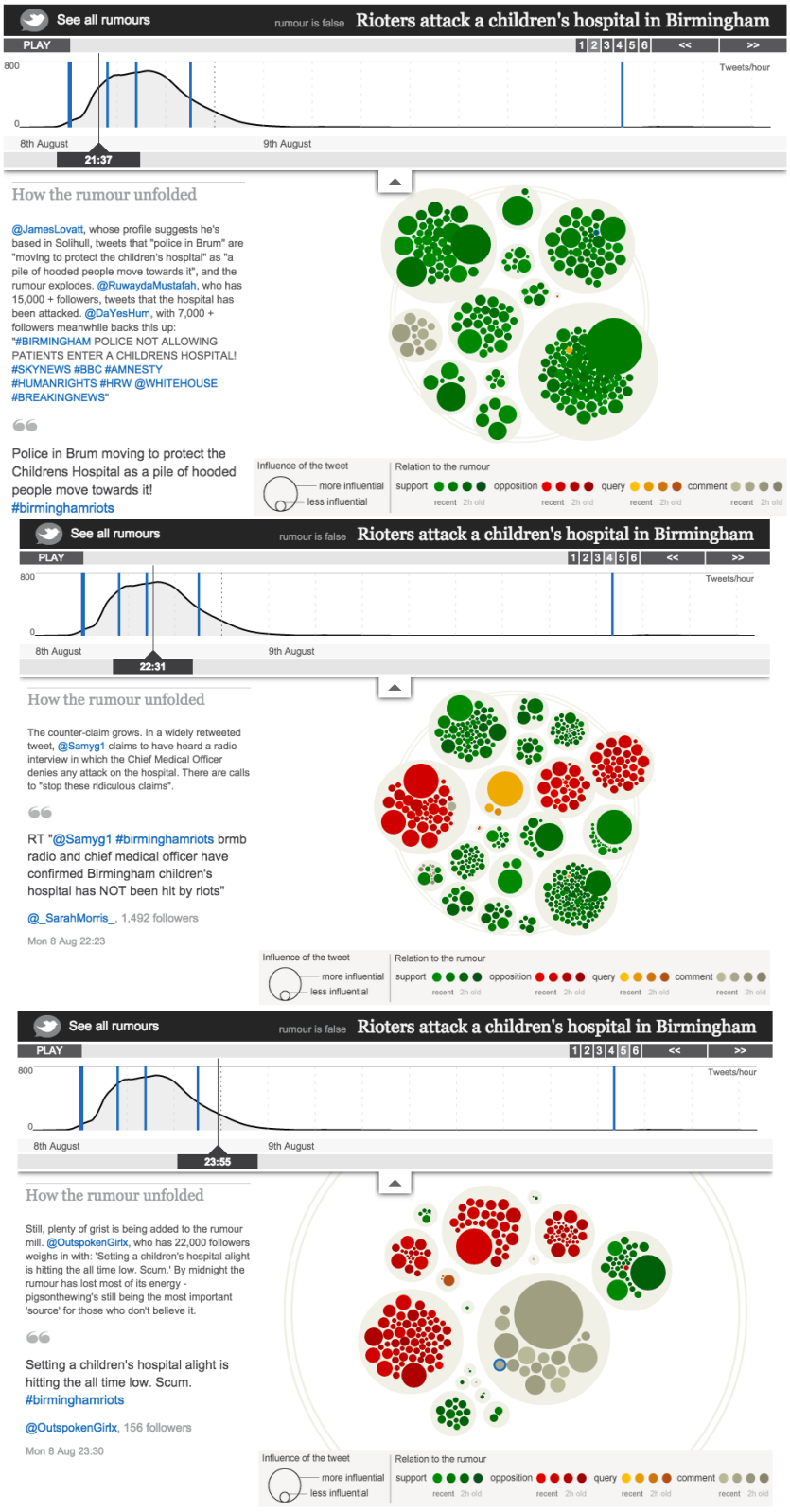


Figure 12: "How Riot Rumours Spread on Twitter", Guardian, December 7, 2011. (a), (b) and (c). Network diagram of clusters of tweets around the rumour "Rioters Attack a Children's Hospital in Birmingham" at three different moments in the unfolding of the rumour and its dismissal.

3.4.5 Revealing Hidden Ties

In this category I grouped narrative views which depict hidden and potentially incriminating sequences of connections or paths between nodes. In this category the nodes represent (alleged) villains as well as their collaborators, whether individuals, businesses or government officials. The edges represent allegedly unethical, fraudulent and potentially incriminating relations. The painting by Marc Lombardi cited in the introduction of this chapter clearly falls in this category.

This category may be interpreted as loosely evoking the network property of "weak ties". Whereas the nodes of a cluster are connected by strong ties, another essential concept in the study of social networks is that of weak ties (Granovetter, 1973). The notion is used to describe connections between nodes belonging to different clusters where nodes through which such connections are established act as bridges. Such nodes have been shown to play an essential role in various social activities through their ability to transport information across the structural holes that separate clusters (Jensen et al., 2015). This narrative view should be distinguished from "exploring associations around single actors" and "detecting key players" in that, while selected actors are at the centre of such narrative views, the emphasis of the reading is not on one or several well-connected actors, but rather on the path of connections that ties the actors and the nature of these ties. I will illustrate this narrative view below with three examples.

The Kansas City Star's "Terrorist Tentacles Know no Boundaries" (November 28, 2004) explores the ties between a global charity and multiple terrorist organisations and supporters, including terrorist Osama bin Laden. In the diagrammatic mode ties or connections are rendered through directed edges represented as arrows, all of which start from a single prominently sized node and point towards several other nodes, which can be construed as individuals or groups based on their icon depictions (see Figure 13). The nodes are further qualified through textual labels which identify them as the incriminated global charity IARA (Islamic African Relief Agency) and the terrorist organisations,

terrorism supporters and terrorists to which IARA is alleged to have ties, including Osama Bin Laden. The incriminating ties are further specified in the headline of the article ("Terrorist Tentacles Know no Boundaries") as well as in the body of the article where the ties are being listed:

> At least eight connections between IARA and Osama bin Laden, his organisations or the Taliban; Two connections to Hamas, the Palestinian terrorist organisations whose suicide bombings ravaged life in Israel; Connections to three other groups that long have been designated as terrorist organizations by federal authorities. (*Kansas City Star*, November 28, 2004)

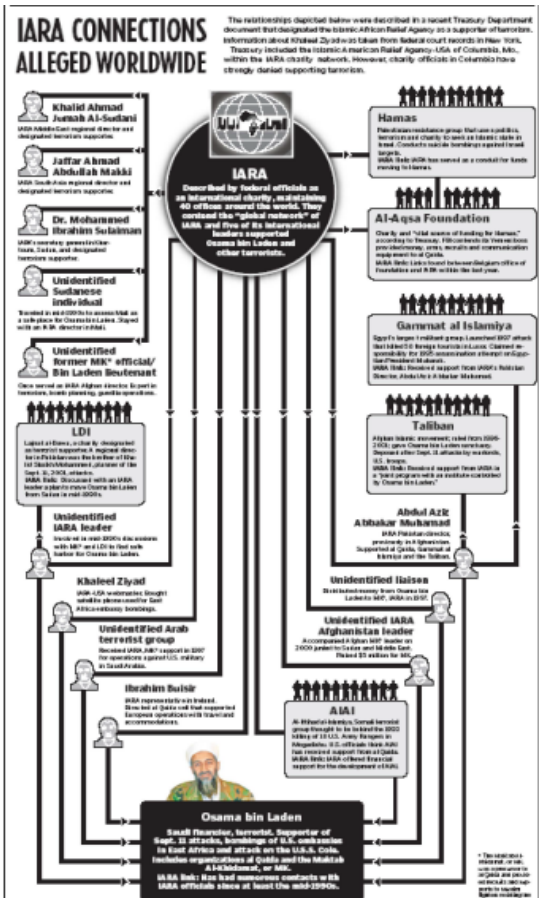


Figure 13: "Terrorist Tentacles Know no Boundaries", Kansas City Star, November 28, 2004. Tree-like network diagram depicting a global charity and its alleged terrorist ties.

Similarly, in *Los Angeles Times*' piece "The Calderon Family's Connections"

(February 21, 2014),[43] which investigates incriminating ties between a powerful Southern Californian family and organisations in its area, both the diagrammatic and the written language modes participate in the narrative reading of incriminating ties. The incriminating ties or connections are rendered through arrows connecting two sets of nodes (see Figure 14). The nodes are qualified through textual labels as the members of the Calderon family and the organisations and businesses with which it has established incriminating relations. The nature of the relations represented by the edges is specified through colour and text: green arrows represent financial donations to campaigns, blue arrows represent legislative interventions or attempts thereof, and yellow arrows represent consultancy services. The headline of the piece ("The Calderon Family's Connections") further anchors this narrative reading. The subtitle of the article qualifies the ties as incriminating through the specification that the Calderon family is under investigation by the FBI. The nature of the ties is also cued by the body of the article, which provides further detail on the investigated connections between the family and various private and public actors in their region.

43 Accessible at: http://graphics.latimes.com/calderon/

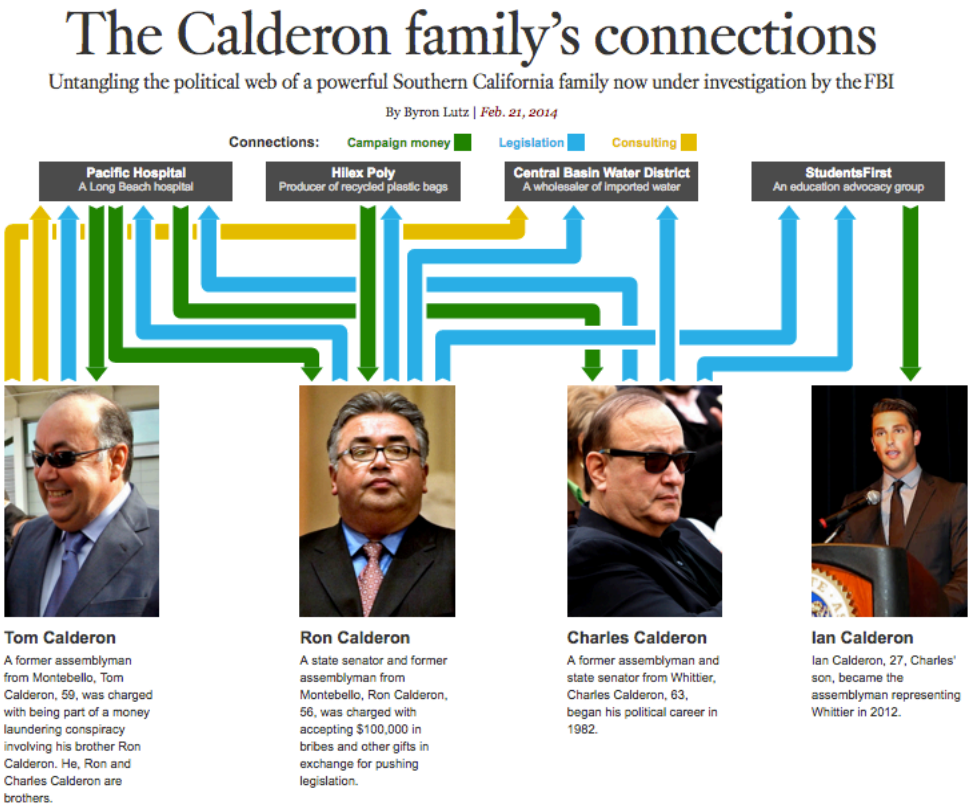


Figure 14: "The Calderon Family's Connections", Los Angeles Times, February 21, 2014. Title, subtitle and interactive network graphic opening the journalistic piece.

Finally, I discuss the construal of incriminating ties in Organised Crime and Corruption Reporting Project (OCCRP)'s interactive piece "The Proxy Platform", published on November 22, 2011.[44] In this piece too, the path is cued by the dynamic diagram and written languages modes. An interactive feature enables the reader to select a node and to visually highlight the sequence of paths or connections to other nodes, as well as to reveal the name labels of these nodes, while the other elements of the graph are dimmed. The sequence of paths is composed of four different types of nodes. The sequencing of the node types on a horizontal axis, inviting (but not restricting) the reading of connections between nodes from left to right further invites the construal of the notion of a path connecting different kinds of nodes (see Figure 15). The node types are specified through labels, from left to right, as: "proxies" (individuals running phantom companies on behalf of the real beneficiaries), "proxy companies" (companies set up to facilitate money

44 Accessible at: https://www.reportingproject.net/proxy/en/

laundering), "banks" through which transactions flow between proxy companies and beneficiaries, and "beneficiary companies" which are on the receiving end of financial transactions). The interpretation of the sequences of connections as incriminating ties would be highly improbable from the diagrammatic mode alone. It is the written language mode, more specifically the body of the article accompanying the interactive graph which describes the mechanisms of the money laundering system which spans Eastern Europe, Central and South America and Asia, that enables the identification of the sequences of paths between nodes as incriminating ties in a money laundering system.

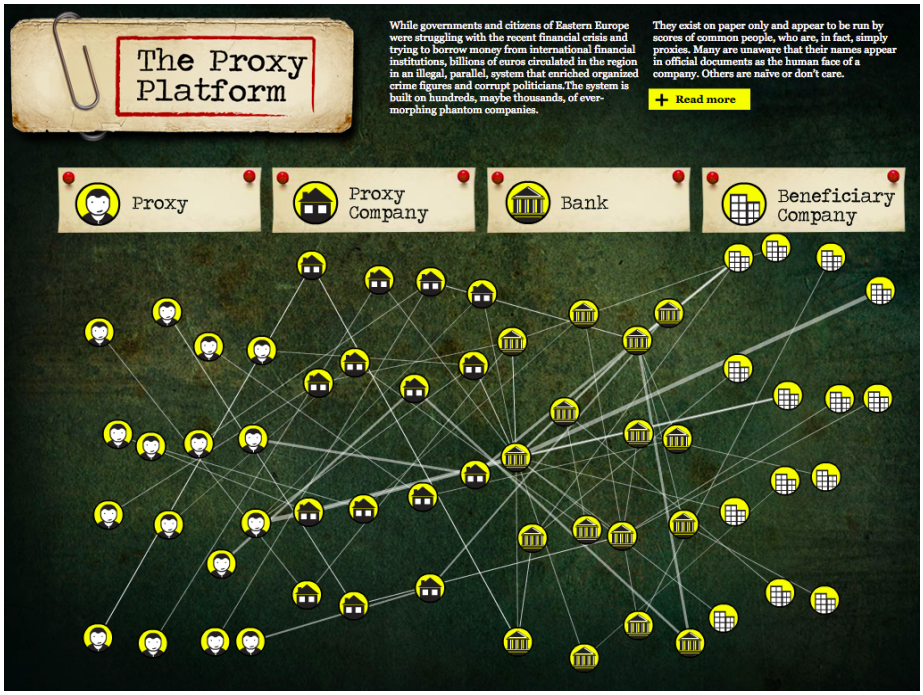


Figure 15: "The Proxy Platform", Organised Crime and Corruption Reporting Project (OCCRP), November 22, 2011. Landing page of the journalistic piece depicting the interactive network graphic of actors that make up the investigated money laundering system.

3.5 Conclusion

In this chapter I took a news device approach to the study of network diagrams and their participation in journalism stories. Following this approach, the focus has been on how this digital object matters or makes a difference to narratives about collective phenomena in the context of journalism projects. I

described its contribution by highlighting five narrative reading patterns that networks elicit in my corpus of journalism pieces.

As colleagues and I discuss elsewhere, common metaphors that associate network graphs with "hairballs" or "spaghetti bowls" point to the absence of established practices for working with the narrative affordances of these objects, as well as to the limited literacy around reading them (Venturini, Bounegru, Jacomy, & Gray, 2017). Although illustrative rather than exhaustive, I hope that this vocabulary of narrative readings as well as the protocol for the construal of narrative meaning out of networks will contribute towards elucidating the meaning-making capacities of these defining visual culture and knowledge-making devices in the digital age, and contribute towards developing literacy to read them narratively.

Methodologically, the contribution of this chapter was to illustrate that existing analytical approaches such as multimodal analysis are well suited to support the analysis of narrative and communicative aspects of digital objects from a news device perspective. Drawing on this approach, I showed that forms of storytelling that rely on network graphics are realised multimodally. It is in the interaction between modes such as static, dynamic or animated diagrams, written language, photographs, layout and pictograms that the specificity of this composite form of storytelling lies.

I show that the narrative readings of networks are co-produced through interactions between the material affordances of networks and the norms and values of the journalistic genres in which they are embedded. First, in keeping with Ryan's (2005) proposition that good narratives think with their medium, that is that they take advantage of the affordances of the medium in which they are realised, I show that the material affordances of these digital objects shape narratives and how collective phenomena become widely knowable. Different narrative views tend to be cued by distinct visual attributes of network diagrams such as node size and position, density of connections, node colour, and size and arrow-like depiction of edges. Narrative views deploy or evoke classic network concepts and properties, from ego-networks, to the power law

distribution, clustering, weak ties and the dynamic character of real-life networks. The production of narrative meaning revolves around relationality or associations and their properties, from topology, to density, to dynamic character or absence. For this reason, journalistic stories that I examined were often configured around the representation and exploration of structures, assemblages or collectives, be they political or power structures and fraud or corruption networks, to mention just a few.

In addition to their resonance with network concepts or properties, I also find these narrative readings to be shaped by some of the values and themes associated with the genres of news and investigative journalism. For example, the narrative readings constructed around the depiction of actors' networks, key players and alliances and oppositions may be interpreted as resonating with news journalism values such as the focus on human interest, personalities, eliteness of news actors, and conflict (Bell, 2007; Dardenne, 2005). Narrative readings constructed around the exposure of incriminating ties resonate with conceptions of investigative journalism narratives to be acting as an instrument for invoking morality in the service of judging civic vice (Ettema & Glasser, 1998).

Returning to the multimodal realisation of these narrative views, besides being cued by particular visual attributes of networks, in all the examined cases the identification and qualification of the actors, connections between them, themes, temporal dimensions and other elements of the narrative reading would have been improbable without anchoring in the written language mode as well as context. The written language and layout modes present in our examples in the form of headlines, article bodies, graphic labels, captions and guides, are so important in the construal of narrative meaning that in this analysis I resolved to eliminate multiple examples of network graphics that had been used in the context of journalism pieces in cases where their original context of publication was unavailable to me and I was thus unable to unambiguously construe narrative meaning. The journalism genre and particularly its convention that the most important information is contained in the headline and lead of an article, and the subjectivity and socio-cultural

knowledge of the reader more broadly, also played an important role in guiding the interpretation of narrative meaning. For this reason, I place emphasis in this analysis on the description of the process by means of which I have *construed* each narrative reading.

Given the small collection of examples I make no claims of comprehensiveness or representativeness of the narrative views which I identified. While for the purposes of this chapter I limited myself to the discussion of five narrative views that exploit or resonate with network concepts or properties that have captured the popular network imagination today, I believe that other narrative readings may also be construed (for example around the concept of "spheres of influence") and invite researchers to conduct further work in this area. To facilitate such work I make available together with this chapter the full collection of examples of journalism pieces deploying network visualisations or concepts and encourage other researchers to expand it.[45]

I also do not claim the boundaries between the identified narrative reading types to be clear-cut. In fact, I encounter in this analysis several borderline cases where, on the basis of the diagrammatic mode alone, the narrative meaning can be equally construed along several narrative views. In these cases, I anchor or qualify the narrative reading by taking into account information present in other elements of the journalistic piece, particularly the headline, lead and body of the article.

I would also like to note that the narrative views which I identified are not mutually exclusive. In fact, in this collection I was often able to construe multiple narrative views from a single journalism piece. These function as building blocks for the broader journalistic narrative. This was particularly the case of pieces drawing on the dynamic diagrammatic mode, where these views were sequences of a larger narrative through which the reader can progress thanks to a strong interactive component realised through a diversity of

45 The list is accessible here: https://figshare.com/articles/List_of_journalism_projects_using_network_concepts_analysis_and_or_diagrams/3126523

techniques and devices. In addition to this, given the aim to qualitatively explore the narrative potential of networks, I have not exhaustively analysed all the narrative views elicited by networks that make up the larger narrative developed in a journalism piece. Instead I selected and limited my focus to illustrating a few narrative views that resonate with classic network concepts and properties. I do however consider the interplay between different narrative sequences in the context of the broader narrative of a single journalism piece to be important and invite researchers to study it.

Finally, to further situate the findings of this chapter, I will address some final broader issues pertaining to the analytical perspective used. I will suggest that while studies that focus on journalistic representations such as the one described in this chapter are necessary and insightful, the frame of study should be extended beyond journalistic representations in their original context of publication to also explore their performativity across different sites and contexts.

In this chapter I focused on the interpretation of these graphics in their narrow context of publication and in doing so I left out both the practices of production of such images, and their circulation and reception dynamics. Such production practices need to be accounted for because visual displays reflect the conventions of the professional practices that create them and the affordances and constraints of the visualisation techniques used, as much as they reflect the characteristics of the depicted phenomena themselves (Bastide, 1990; Lynch, 1985). As far as circulation dynamics are concerned, journalistic images and visualisations may take on different roles, functions and interpretations depending on who they reach and the various situations in which they are received and used.

In the context of the social study of art, Becker (1982) proposes the concept of "art worlds" to emphasise the need to expand the analytical frame beyond the familiar art works themselves to how they are circulated and constructed through "the network of people whose cooperative activity, organised via their joint knowledge of conventional means of doing things, produces the kind of

art works that art world is noted for" (p. 10). In the next chapter I shift my focus from news works to the journalism coding work which increasingly underpins the production of journalistic projects, news tools and infrastructures. In doing so I also shift focus from digital objects as sites of narrative and communicative acts, to digital objects as sites of platformisation of news work (Chapter 4) and commodification of news audiences (Chapter 5).

4. Platformising News Code: GitHub as a Connective Coding Device

In this chapter I interrogate the online platform as a news device, and more specifically the code sharing platform GitHub.[46] Following the news device perspective, I organise the chapter around two questions. The first concerns how coding is structured through GitHub's technical infrastructure in accordance to its economic imperatives, and, the second one pertains to, exploring the platform-specific ways of studying coding that GitHub affords.

According to Srnicek (2016), online platforms can be understood as "digital infrastructures that enable two or more groups to interact" (p. 25). I focus on GitHub for a number of reasons. GitHub is the largest online software development and code hosting service (Gousios, Vasilescu, Serebrenik, & Zaidman, 2014), and the most used one by news organisations (Usher, 2016). The platform too recognises news organisations as one of its stakeholders by prominently featuring a curated list of "open journalism" repositories in its collections section, which showcases work from industries whose presence is rapidly growing on GitHub.[47] Studying GitHub is also relevant in relation to a growing area of interest in journalism research concerned with the study of coding, computation and open source software development (see, e.g., Lewis & Usher, 2013).

Online platforms more generally have become important actors in today's news media (Nielsen & Ganter, 2018). When I began this study, GitHub was not part of GAFAM, the group of largest and most influential digital platforms on which digital cultural production in the Western societies is dependent, and

---

46 This chapter is informed by a pilot study I co-led with Jonathan Gray and Stefania Milan at the Digital Methods Summer School 2015 at the University of Amsterdam. The project team included: Jonathan Albright, Matteo Azzi, Stefan Baack, Stefano Bandera, Rishabh Dara, Rebeca Diez, Sylvain Firer-Blaess, Ivo Furman, Robert Gutounig, Janna Joceli, Cristel Kolopaking, Lisa Krieg, Lisa Langenkamp, Sam Leon, Sjoukje van der Meulen, Mariola Pagán, Tamara Pinos, Ana Pop Stefanija, Tim Riley, Richard Rogers and Savaş Yıldırım. While this chapter does not document the outcomes of the pilot study, this early collaborative phase was useful in testing research design and methods. This phase also benefitted the researchers and students involved as they were exposed to collaborative working formats and new methods and research objects in a guided environment. This is consonant with the approach of this thesis not just to provide better accounts of social lives of digital devices in journalism, but to provide perspectives and methods which could be used by and useful for others.

47 See https://github.com/collections/open-journalism

which includes Google, Apple, Facebook, Amazon and Microsoft (Nieborg & Poell, 2018). With its acquisition by Microsoft in 2018, GitHub has joined this group.

These platforms are increasingly influencing the way news is produced, distributed and monetised by reshaping the news media infrastructures that underpin these processes (Nechushtai, 2018; for the way in which online platforms are reshaping web infrastructures on which news media also rely, see, Helmond, 2015a). While GitHub may not have the influence that Facebook and Twitter have over news processes and infrastructures, the platform is interesting to examine through the news device lens not because of the scale of its influence but because of its particularities. According to Mackenzie (2018), as code increasingly underpins many of the infrastructures that make up collective life and professional domains today, the particularity of GitHub lies in its growing role as a platform for building infrastructures. Indeed GitHub hosts some of the code that underpins platforms such as Facebook and Twitter, and also some of the source code for news infrastructures such as *The Guardian* website.[48] As GitHub is increasingly used by newsrooms, the platform can be seen as one of the sites where digital news infrastructures are configured, assembled, maintained, and sometimes made public, and where all these activities and processes are recorded and documented.

In this chapter I explore how GitHub participates in the configuration of journalism coding work and news infrastructure making, and how it can become a resource for studying the participation of news initiatives on the platform. By attending to how the platform intervenes both in journalism coding work and in approaches to studying it, I aim to illustrate the two aspects of the news device approach.

The chapter is organised as follows. First, I introduce how the relationship between GitHub and journalism has been discussed so far, and some of the ways in which GitHub has been studied so far. Next, I examine how GitHub

[48] https://github.com/guardian/frontend

configures code development from a material-economic perspective. And finally, I examine how networked code repositories can be repurposed as a research device to examine how news initiatives participate on the platform.

4.1 From User to Platform Studies Approaches to GitHub

In the context of journalism, GitHub has been linked to open source software development practices. In *Interactive Journalism: Hackers, Data and Code*, Usher (2016) documents the encounter between code sharing platforms and journalism. She links the rise of open source software in news to the development of code sharing platforms such as SourceForge, Google Code and GitHub. While the open source movement has a longer history, Usher (2016) argues that for newsroom journalists "the rise of open source didn't really make much difference until the early 2000s, when platforms were developed to host open-source code in a way that helped promote social networking" ("Social Open Source", para. 1). She refers to the intertwinement of open source coding with social networking platforms as "social open source". According to Usher (2016), while both SourceForge and Google Code were previously used by news organisations, now most of these use GitHub.

GitHub has also been linked to effort to make news work public and facilitate greater transparency and accountability (Keegan, 2014; Stark & Diakopoulos, 2016; Usher, 2016), particularly in relation to the use of code, data, software, algorithms and computation, in the context of computational journalism (Hamilton & Turner, 2009; Flew, Spurgeon, Daniel, & Swift, 2012), data journalism (Coddington, 2015; Gray et al., 2012), algorithmic accountability reporting (Diakopoulos, 2013, 2015) and automated journalism (Graefe, 2016). In the context of interactive journalism, Usher (2016) argues that GitHub "as an open sharing platform, allows interactive journalists (and the public) to see the backbone of what's underneath these journalists' efforts. The news work, albeit written in code, is unmasked" ("Expressions of Openness through Journalistic Work", para. 4).

The role of social network functionalities (also known as "social coding") in collaborative code development, and the activity transparency that GitHub enables, are also key themes in the extensive academic literature about the platform. Particularly in computer and information science and software engineering, GitHub has become an object of study in relation to collaborative software development. Studies examine the impact of social network functionalities on peer and code assessment and the presence of emotions and network structures in collaborative software development (see, e.g., Dabbish, Stuart, Tsay, & Herbsleb, 2012; Guzman, Azócar, & Li, 2014; Marlow, Dabbish, & Herbsleb, 2013; McDonald & Goggins, 2013; Thung, Bissyande, Lo, & Jiang, 2013; Tsay, Dabbish, & Herbsleb, 2014). They also examine productivity and quality in software development (see, e.g., Ray, Posnett, Filkov, & Devanbu, 2014; Vasilescu, Posnett, et al., 2015; Vasilescu, Yu, Wang, Devanbu, & Filkov, 2015). In these same disciplines GitHub is also discussed as a source of data to be mined and the problems therein (see, e.g., Gousios et al., 2014; Kalliamvakou et al., 2014; Russell, 2013).

Much of the literature discussed so far is in various ways concerned with what Bucher (2012) would call "usage studies", i.e. a focus on understanding motivations and various aspects of practice from the perspective of the user and less so with the critical investigation of the platform itself as the interplays between technical infrastructure and economic imperatives that underpin platform use. A notable exception in this sense is "Metacommunities of the Code Sharing Commons", an exemplary study of code sharing practices on GitHub involving Adrian Mackenzie, Matthew Fuller, Andrew Goffey, Richard Mills and Stu Sharples.[49]

Taking a platform studies perspective, Mackenzie (2018) argues that the study of platforms should account not just for how they are used but also for how they structure social practices through processes of platformisation and capitalisation, both of which are central to the making of platforms. This is because platforms are not "stable backgrounds" for social action (Parikka,

49 http://metacommunities.github.io/metacommunities/

2011), but are constantly being (re)configured. Platformising processes render "an ensemble as a platform" (Mackenzie, 2018, p. 37). They do so by "engineer[ing] specific forms of connection" and "placing people and things into specific kinds of relations", according to the platform's capitalisation or economic imperatives (Mackenzie, 2018, p. 37). Capitalisation processes enable the "conversion of hitherto intangible or uncontrolled social processes into potential asset streams", by configuring a particular setting into a site that has the potential to generate "a steady stream of future cash flows" or capital (Mackenzie, 2018, p. 37).

Drawing on Mackenzie (2018) and on platform studies as discussed in Chapter 2, in what follows I examine how GitHub structures journalism coding.

4.2 From Social Coding to Connective Coding

To understand how GitHub participates in the configuration of journalism coding and the making of news infrastructures, I will briefly discuss some of the platformising processes that underpin it. This is based on a careful examination of GitHub's various interfaces, including its user interface and API, as well as platform documentation from help pages, the platform's development blog and technology press articles. Drawing on van Dijck (2013), I will argue that through these processes GitHub renders coding practices on the platform not simply as social coding but as connective coding.

4.2.1 GitHub's Multiple Stakeholder Market

Founded in 2008, GitHub is a for-profit private company. The GitHub platform is based on Git, an open source distributed version control and source code management system. This system was created in 2005 by Linus Torvalds to support collaboration on the development of the Linux operating system. GitHub enhances Git with web-based hosting, browser-based and desktop-based graphical interface and social networking functionalities. GitHub functioned on its own revenue in the first years of existence and took

venture capital four years after its funding, in 2012 (Gannes, 2012). In 2018 the company was bought by Microsoft. GitHub positions itself as the "world's leading software development platform" (n.d.-a), and previously it referred to itself as a "social coding" platform and a "code sharing platform". GitHub also refers to itself as the "largest open source community in the world" (n.d.-j), a frame which is often taken up by the press.

References to GitHub as an open source code repository and as a code sharing platform co-exist alongside business reports that document GitHub's financial position. According to such reports, around half of GitHub's 2017 annual revenue was generated by businesses maintaining private repositories and the other half by individual users or teams using private repositories (Petersen, 2017). The GitHub Octoverse 2017 report states that around half of the largest US based companies use GitHub services, including Microsoft, Facebook and Google (GitHub, 2017).

The facilitation of these seemingly divergent use cases and groups of end-users is not an anomaly but a constitutive and defining feature of online platforms. This is because online platforms act as technical infrastructures that negotiate and enable interactions between two or more different groups of end-users or stakeholders and between heterogenous goals and interests (Gillespie, 2010; Srnicek, 2016). In the case of GitHub, these include open source and individual developers, businesses, and third-party application and service developers.

From an economic perspective, platforms such as GitHub are understood as "multi-sided markets" (Evans, Hagiu, & Schmalensee, 2006; Rochet & Tirole, 2003). According to Rieder and Sire (2014):

> A multi-sided market consists of a platform that brings together at least two distinct groups of end-users. The attendance of end-users on one side creates a positive externality which makes the good sold on the other(s) more attractive, and vice versa. A platform that enables interactions between the parties can internalize this kind of externality – that is, make profits." (p. 199)

As a "product platform" (Srnicek, 2016) which generates revenue by providing a service against a subscription fee, GitHub is organised around the freemium model, whereby free repository hosting is provided for public projects and paid hosting for private repositories. It is thus businesses and individuals who code privately that enable GitHub to provide free hosting for public and open source software and to be at once the largest public code repository in the world and draw its revenues entirely from private hosting of software development projects. In this model GitHub subsidises open source developers, or rather individuals and organisations who code publicly,[50] and charges businesses and developers who code privately, by providing various paid plans for private repositories. The two models might not target different user groups but different use cases, as organisations might use private repositories for everyday work and release code publicly using the public repository function, as Fuller, Goffey, Mackenzie, Mills and Sharples (2017) suggest. Indeed, GitHub's annual public reports show that large numbers of business developers contribute to open source projects on GitHub (GitHub, 2017; see also Ghoshal, 2016). Moreover, developers and organisations who host their code publicly are not only providing a vast collection of open source projects that businesses can draw on, but they are also seen as marketing and conversion tools to attract enterprise solutions, as developers who make personal use of the platform become hired by businesses (Sullivan, 2016). GitHub also charges third-party app developers who build apps to enhance GitHub's functionalities and advertise them on the GitHub Marketplace (through a 25% share from any app sold on its marketplace). In order to support the use of platform data to develop a third-party ecosystem of products and services that enhance the platform experience, GitHub does not charge for API usage that complies with its rate limits. It does however charge for API usage that goes beyond its rate limits or that results in the marketisation of services that mimic the GitHub service experience (GitHub, n.d.-i).

50 As GitHub does not enforce licensing for projects it hosts, public projects may or may not be open source. By default if a public repository does not have a license specified then it will be considered an all rights reserved repository (Phipps, 2012).

4.2.2 Making Participation on the Platform Economically Valuable

An essential platformising process that enables the facilitation of these heterogenous groups and interests into the marketplace configuration described above, is what van Dijck (2013) calls "turning connectedness into connectivity by means of coding technologies" (p. 16).

While developers and organisations engaging in public or open source coding (such as the journalism initiatives studied in this chapter), do not pay fees for their use of the platform, they do however contribute the publicity, traceability, metrification, analysability and valorisation of their participation on the platform.

Conditions for participation are set through GitHub's technical infrastructure in alignment with the platform's economic aims. This includes a front-end which seeks to solicit, intensify and accelerate user engagement and a back-end comprised of servers and data storage, mining and archival capabilities (Gehl, 2011). These features, just as in the case of other social media platforms, are organised around nurturing a platform ecosystem that would multiply valorisation of connectivity around several registers (Gerlitz, 2016; Marres, 2017a).

More specifically, following the model of other social media platforms, participation in software development is made economically valuable by setting up an infrastructure that pre-defines and standardises possibilities for user action and the forms that these take at the user interface level (Gerlitz & Rieder, 2018; Gray, Gerlitz, & Bounegru, 2018). In the case of GitHub this would include enabling forms of action such as "committing", submitting "pull requests", "forking", "starring" or "watching" repositories, and following users. Through each of these actions, connections are recorded between users, and between users and objects such as repositories. In addition to this, free repositories are subjected to an "environment of information transparency" (van der Vlist, 2013, p. 20) whereby every activity associated with them is attributed and visible to anyone who accesses the platform (see Figure 16).

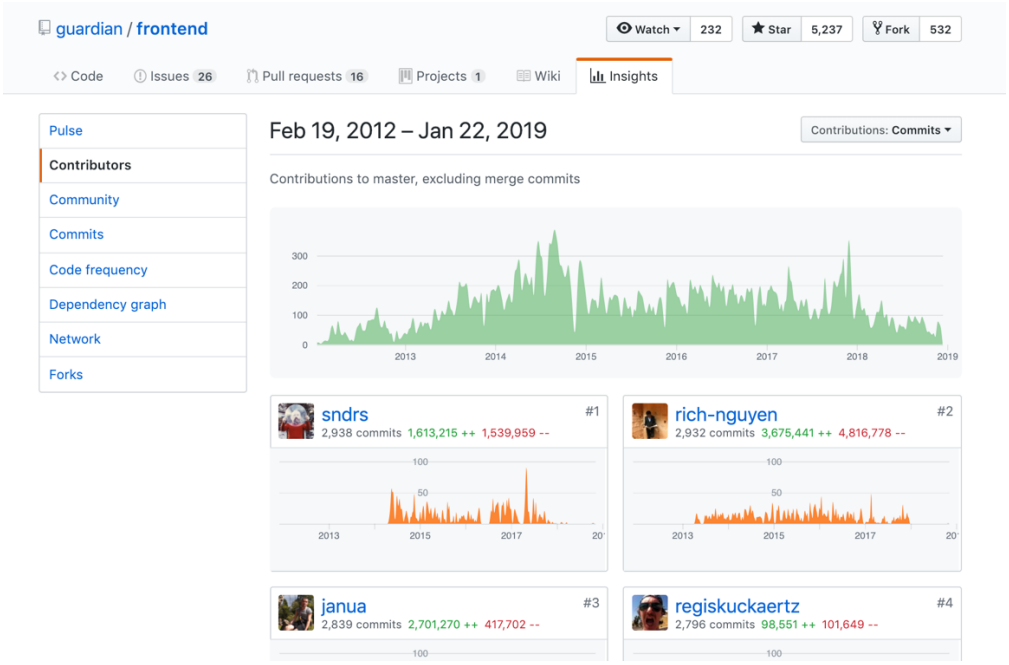


Figure 16: Screenshot of browser-based user interface view of public GitHub repository showing the top contributors to a public repository based on their number of commits. In the top right corner the social counters associated with repositories can be noticed.

Standardising possibilities for action enables social media platforms like GitHub to render selected activities, projects and people measurable, calculable and comparable (Gerlitz & Rieder, 2018). This is done through social metrics and counters such as GitHub's "forks count" and "stargazers count" (visible in Figure 16), rankings such as its "trending" feed, and other calculations and statistics released by the platform, such as through the annual "State of the Octoverse" report (GitHub, 2017).

By making projects and people commensurable through the introduction of common metrics, the platform materialises an auditorial culture (Gane, 2014; Power, 1994, 1999, 2000; Strathern, 2000) based on quantitative measures, that intensifies evaluation and competition between projects:

> The units made available through the above-mentioned platforms are vying for attention and the omnipresence of counters and rankings that measure and evaluate popularity, relevance, or authority is essential to reinforcing a state of permanent competition where everyone is

> constantly updated on where they stand. (Rieder, 2017, p. 6)

This environment is generative of reactivity dynamics (Espeland & Sauder, 2007) whereby users modify their behaviours in response to the evaluations that the platform makes available. As platforms are organised around "the pursuit of participation" (Bucher, 2012b, p. 10), this is a dynamic that platforms welcome and encourage (Gerlitz & Lury, 2014). The constant pressure to perform to the metrics and gain visibility, is generative of both grassroots responses in the form of how-to guides for increasing visibility by accumulating currency, i.e. stars (see, e.g., Neculai, 2017; Suri, 2017), as well as commercial initiatives that enable the purchase of false stars and followers (Governor, 2017).

Moreover, the metrification of coding and engagement acts on the platform is accompanied by an intensification and multiplication of social dynamics (Bucher, 2012b; Gerlitz & Helmond, 2013). For example, GitHub displays user activities on the user's home page, as well as on the news feeds of her followees. Activity around a repository that a user watches or subscribes to is also displayed in the user's "news feed". Repositories are recommended for starring and bookmarking based on the user's own activity on the platform.

All these data streams and recommendation features aim to intensify activity on the platform by inviting further engagement and suggesting possibilities for action. In doing so they multiply the collection of economically valuable user data (Gerlitz & Helmond, 2013) and the growth of the platform's connective assets, i.e. products that result from the conversion of social coding into economically valuable connective coding. These connective assets may include platform data and knowledge products, which in turn become part of the platform's offerings to its stakeholders (Gerlitz, 2016; Rieder, 2017). As described above, these may be embedded in various interfaces, e.g. as recommendations and trends via the browser-based user interface, as platform data via the application programming interface (API), and annual reports about platform accomplishments geared towards attracting investors and clients.

The capture of social dynamics and behaviours around public repositories and

the creation of knowledge or information products is not only essential to GitHub's functioning as a marketplace, but also enables GitHub to function as an ordering mechanism (Rieder, 2017) for the public and open source coding space. This has social consequences in terms of how attention is guided to particular repositories and people (Rieder, 2017), what is privileged and what is devalued, as well as for software development and infrastructure making practices.

4.2.3 Connective Coding

Having briefly described some of the platformising processes at work in configuring GitHub as an online platform, I will now elaborate on the implications of these processes on coding practices and infrastructure making, which I capture with the notion of *connective coding*. The notion of connective coding moves away from seeing GitHub just as a container for software code and a management system for software development, to pay attention to the infrastructures and processes through which coding is platformised. The notion is intended to expand the understanding of social coding beyond the connectedness that GitHub enables to also capture the commodification of social coding by "turning connectedness into connectivity by means of coding technologies" (van Dijck, 2013, p. 16). Mosco (2009) defines commodification as "the process of turning use values into exchange values, of transforming products whose value is determined by their ability to meet individual and social needs into products whose value is set by their market price" (p. 132). In the context of GitHub, the commodification of social coding refers to the potentiality of accumulation of economic capital by converting public coding activities, developer profiles and behaviours into assets that may attract future revenue and investment to the platform (Mackenzie, 2018), through the platform's technical infrastructures. These connective assets have the potential to be capitalised by platform as well as by third parties that make up the platform ecosystem (under the terms set by the platform), and to enter various economies and forms of valuation (Gerlitz, 2016).

For example, numbers, of users and repositories accumulated by the platform, play an important part in enacting platform growth and in supporting future investment of capital (Mackenzie, 2018). The release of data points via an API enables third parties to derive their own forms of value, be it economic, cultural or social (Gerlitz, 2016). In addition to being used to improve the performance of the platform according to its aims, and to optimise software development processes, due to its large base of public developer profiles and code repositories, GitHub has become a recruitment tool for the technology industry and has come to be seen as the best resume for a developer, displaying markers of reputation, productivity and uptake (Fuller et al., 2017; Petersen, 2017). GitHub is also becoming a key provider of data about the software development sector, supporting a number of startups that provide various kinds of data mining, analytics and recruitment services in this sector (Richtel, 2013). In addition to this, as touched on in the previous section, a large corpus of research into collaborative and open source software development also increasingly relies on behavioural data collected and released by GitHub.

Platformised coding also draws attention to how coding and engagement work also become a form of platform work, in the sense that every coding act also contributes to enacting the platform ensemble. This is illustrative of the asymmetries that characterise online platforms, between the actors who define "conditions for participation" (Gerlitz, 2016, p. 19) by setting up a techno-commercial infrastructure (the platform owner), and platform users, as well as between the capacities of datafied users and those of third-parties which derive economic value from such datafied user activities (Gerlitz & Helmond, 2013; Lehtiniemi, 2016), such as startups and app developers.

Finally, connective coding plays out through the "double articulation" (Langlois, Elmer, McKelvey, & Devereaux, 2009, p. 415) of platform infrastructure and social practices. An example of how social practices may contribute to the articulation of platforms is the case of users deploying different kinds of tactics to negotiate the conditions of participation set by the platform, In the case of GitHub, in 2016, users frustrated with various issues pertaining to the technical configuration of GitHub and its use affordances,

produced a critical response in the form of letter addressed to platform owners and hosted on the platform, [51] as a way to prompt a response and reconfiguration of interface features to better support public coding work.

Whereas the notion of connective coding draws attention to how particular platformising processes may configure social practices unfolding on the platform, they do not over-determine them. Just as platforms are not fixed or stable arrangements, social practices are also not uniform materials to be platformised but fluctuating, diverse and variable. Furthermore, platforms and social practices are not separate entities but mutually articulating (van Dijck, 2013). For this reason, it is necessary to examine how connective coding plays out in issue or domain specific contexts. In the next section I examine how GitHub is mobilised for journalism coding and what platform-specific ways of studying coding GitHub affords.

4.3 Exploring Journalism Coding on GitHub

Following the news device perspective, in this section I explore how connective coding can be repurposed as a research device, by "seek[ing] to derive … analytic capacities from the pre-formatting that is distinctive of digital social and cultural data" (Weltevrede, 2016, p. 42). More specifically, I explore the ways in which repositories and coding acts are formatted, counted, ranked and how these can be configured to study specificities, modulations and variability of platform-specific news coding practices.

I focus on three analyses to illustrate aspects of connective journalism code dynamics on GitHub. The first analysis repurposes GitHub's forking feature to draw attention to connective journalism coding as imitative work, the second one repurposes collaboration and engagement metrics to explore how journalism code repositories are ranked and valued, and the third one repurposes date stamps associated with repositories to explore the temporal dynamics of journalism code. Before describing the outcomes of these analyses

[51] https://github.com/dear-github/dear-github

I describe my research process.

4.3.1 The Research Affordances of Platformised Repositories

According to Weltevrede & Borra (2016), platform data is the outcome of intersecting socio-technical processes at work in platform-making: interactions between users and platform, the preformatting and curation of user action through the platform infrastructure, how platform affordances are activated through use practices, and how platform activity is captured and recursively delivered back to various platform stakeholders in various forms, as social counters, via APIs, news feeds, trends or recommendations.

One type of platform data that GitHub generates revolves around its repositories. Journalism code inhabits GitHub in the form of repositories, associated with GitHub user or organisation accounts. As one of GitHub's "primary asset form[s]" (Mackenzie, 2018, p. 45), repositories are "named collections of files mainly containing code, but also a great variety of operational documents (settings, manuals, installation instructions, etc.), stored in many versions and varieties" (Mackenzie, 2018, p. 40). Repositories may be private or public, depending on the user's chosen pricing structure. Whether a repository is public or private has important consequences for the way in which it is platformised and hence for its research (or analytical) affordances.

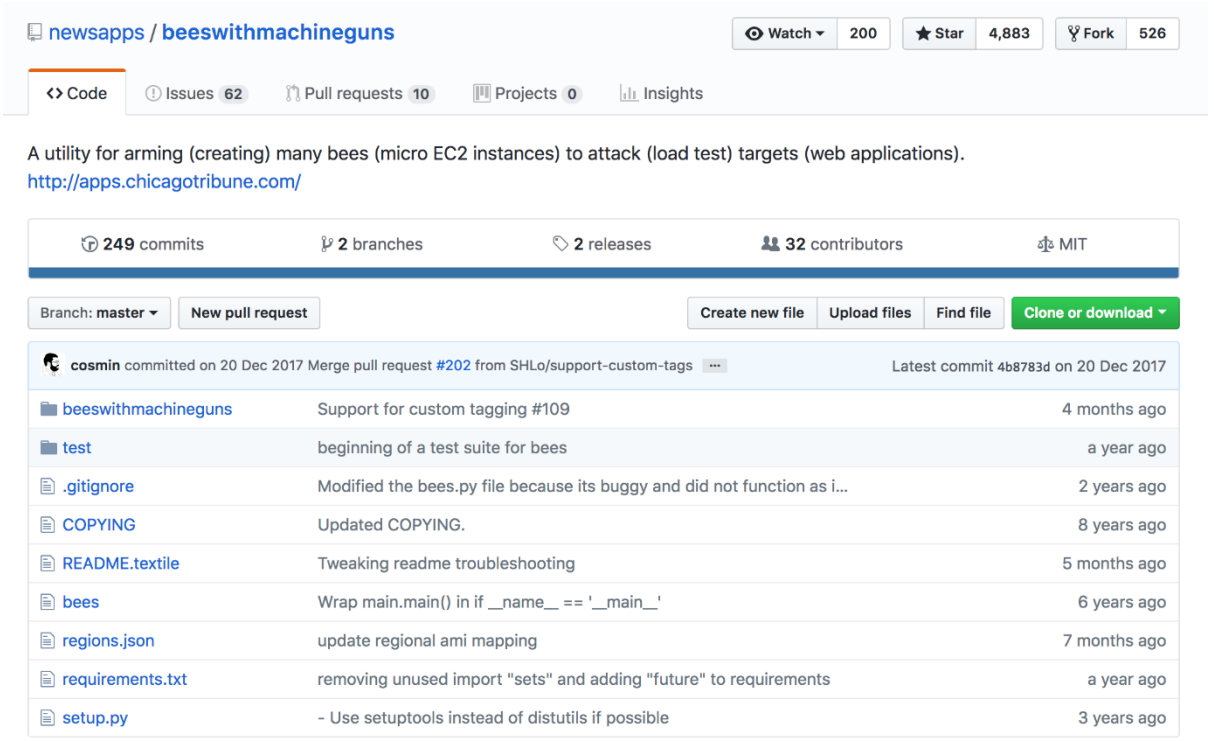


Figure 17: Screenshot of browser-based user interface view of public GitHub repository. GitHub compares repositories with folders where users can store their project files together with their revision histories.[52]

First, public repositories' content and metadata are by default subjected to data mining by the platform. This operation is necessary and underpins the connectivity and calculative operations undertaken by the platform such as the trending feature, and the delivery of recommendations, including repository security recommendations (GitHub, n.d.-c). Public repository metadata is not only data-mined by the platform for its day-to-day operations but is also made available for integration in third-party products and services via the GitHub API (GitHub, n.d.-g). On the other hand, private repositories are not data-mined by default. Owners of private repositories have the option to opt-into GitHub's tracking of their repository metadata, whereas public repositories do not have the option to opt-out of platform tracking (GitHub, n.d.-c).

Through this configuration, the platform guides researchers towards public repositories as an object of study. But it is not only the degree of accessibility that drives researchers towards public repositories. Public repositories are a

[52] https://help.github.com/en/articles/about-repositories

more prolific site for studying *connective* coding. This is because they are more intensely subjected to platformising processes such as those described in the previous section (e.g. metricisation, datafication, recommendation and ranking), as they are configured as connective assets that are essential for attracting a number of different stakeholders to the platform.

4.3.2 Demarcating the Journalism Code Space Through an Expert List

To illustrate how public journalism coding on GitHub can be examined from a news device perspective, I use a collection of 3,665 journalism code repositories. While Fuller et al. (2017) argue that the manageable size of the platform encourages the study of "the whole of GitHub", as the researchers themselves admit, a big data or "n=all" approach obscures as much as it provides insights in the case of GitHub. One of the things that is obscured, as the researchers found, is precisely the diversity of practices that inhabit the platform and the specificities of situated, domain-specific practices (Mackenzie, Fuller, Goffey, Mills, & Sharples, 2014).

For this reason, I focus on a limited collection of journalism repositories which is derived from an expert list of public GitHub organisation accounts associated with journalism coding. The list is maintained by the specialty publication *Source* and is largely Anglo-American in focus. *Source* is part of OpenNews, a programme of the Knight and Mozilla Foundations until 2017. The programme was set up to facilitate an open source ethos in newsrooms (Lichterman, 2017) but recently the publication distanced itself from open source as its primary goal in order to more smoothly align with the needs of journalism coders. Now the publication emphasises a focus on "journalism code and the community of developers, designers, journalists, and editors who make it" (Source, n.d.).

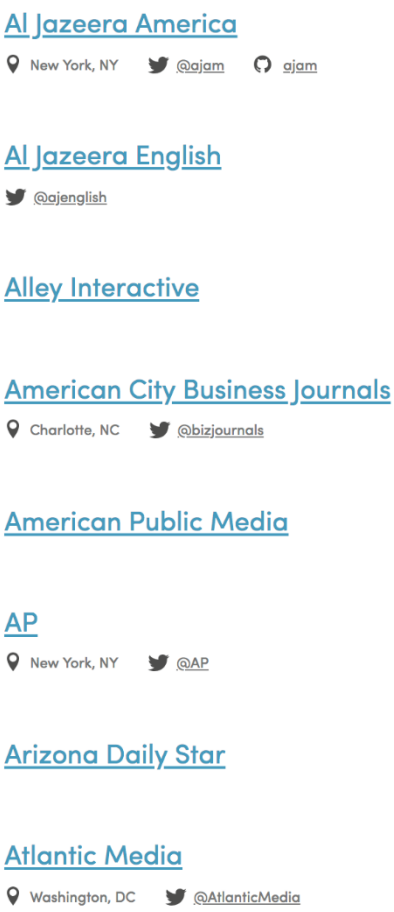


Figure 18: Screenshot of Source's listing of organisations engaging in journalism coding, with links to their GitHub and Twitter accounts.

Another expert list is maintained by GitHub itself, under its "open journalism" collection, but it lists only fifteen repositories.[53] In recent years, a couple of grassroots lists maintained by newsrooms and journalists also emerged, to which I am also contributing.[54] However, at the time when the analyses for this chapter were conducted in June-July 2016, the *Source* list was one of the most prominent and comprehensive lists available, which is why I chose it to illustrate this approach. *Source* also maintains a list of developers engaged in journalism coding. But for the purposes of this chapter I focus on GitHub organisations as organisations have been found to drive activity on GitHub (Mackenzie et al., 2014). A script was used to scrape the GitHub accounts of

53 https://github.com/collections/open-journalism
54 See https://github.com/MinnPost/car-code and https://github.com/silva-shih/open-journalism.

these organisations from the list maintained by Source.[55] Future research should be extended to the grassroots lists available today to enable comparison, as well as to the list of programmer-journalist accounts.

To generate the collection of public repositories of journalism code used for this study, the expert list was filtered to remove GitHub accounts that were associated with organisations or initiatives that were not primarily journalistic in focus, such as Twitter's GitHub account. Missing GitHub accounts of listed organisations were manually added, resulting in a list of 87 public GitHub organisation accounts, be they newsrooms or other journalism initiatives, such as collaborations between journalists around particular projects. The 3,665 code repositories associated with these 87 accounts constitute my research corpus.

4.3.3 API Calling as Data Collection Technique

To collect data about these repositories I use an increasingly common technique in digital social research and outside it, known as API calling (Weltevrede, 2016; Rogers, 2017). API calling is an automated technique of data retrieval from APIs. APIs are "protocological software objects" that "allow interested parties to access the data and functionality of popular online services, all in a very controlled manner" (Bucher, 2013, p.1). As discussed in the first section, APIs are part of the business model of social media platforms, enabling interoperability between the platform and other software systems which enable the platform to extend beyond itself (Bucher, 2013), through integrations and applications.

APIs do not allow access to all platform data but to selected data-points and are subject to particular conditions of access, from authentication to rate limits and data access limitations. Moreover, with APIs, data collection is organised as an "operation of extraction from the negotiated interests and cultures of

[55] The script was written by Sam Leon, data analyst at Global Witness, and is available at: https://github.com/noelmas/dmi-summer-scrapes.

use" of the platform (Weltevrede, 2016, p. 34). Organising data collection through API calls is illustrative of the "distributed" character of digital social research (Marres, 2012b), and the participation of digital devices not just in news work but also in journalism research.

The GitHub API provides different end points that enable the collection of data about repositories. For example, the Events API enables the collection of public events associated with repositories in close to real-time, but significantly limits access to past events by allowing capture of events no older than 90 days. In addition to this, the API is subject to rate limits (300 events per call) and call frequency limits (GitHub, n.d.-b). There is also a GitHub Archive[56] that records the public GitHub timeline but this is limited to events taking place between 2011 and 2015. For these reasons, in this study I decided to use the repository end point. The repository end point of the GitHub API allows developers to programmatically alter, extract data and make changes to repositories. The repositories maintained by these journalism code organisations and their associated metadata were extracted in June-July 2016 via the repository end-point of the GitHub API through a number of scripts co-developed with the Digital Methods Initiative (Digital Methods Initiative, n.d.-a). The scripts have interfaces which allow users to input lists of GitHub organisation, users and repositories names, and extract their associated metadata from the relevant end points of the API.

The result of the extraction process is social transaction data for 3,665 journalism code repositories, including repository creation and update date stamps, popularity and collaboration metrics. The configuration of this data, through pre-formatting of user actions at the user interface level and at the API level, conditions analytical possibilities. Weltevrede (2016) suggests that importing empirical categories that come with using pre-formatted platform data can be seen as an enhancement of the analytical capacities of digital research, as it enables the researcher to study use practices but to also include "the operations of digital media into analysis" (p. 34). When applied to the study of news, such approaches enable us to make findings not only about

56 https://www.gharchive.org/

news and journalism but also about digital devices themselves. This is because researchers need to account for how the platform intervenes in the object of study. Hence, examining how journalism initiatives inhabit GitHub is also a way of exploring the platform, given that inhabiting the platform is an interaction between user and platform features.

However, while the device perspective encourages to "follow the medium" (Rogers, 2013) in the collection of data and composition of analytical categories, the researcher nevertheless needs to work to align these with her research interests and questions, which might sometimes involve pushing against the analytical directions inscribed in platform data and carving forms of analysis that are better aligned with the research objectives at hand (Marres & Gerlitz, 2015; Venturini et al., 2018). All these research decisions will be illustrated in the analyses below.

4.3.4 Imitation and Originality in Journalism Coding Work

A specificity of public coding on GitHub is that through its technical infrastructure based on Git, GitHub institutes copying of code as a user right through its fork function (GitHub, n.d.-h). By automating and simplifying the copying of repositories through the fork button, GitHub institutes a mode of production based on imitation and variation (Fuller et al., 2017), which is one of the ways through which the platform seeks to intensify relations between its stakeholders.

The automation of acts of imitation through the fork button holds analytical capacities. Indeed, several studies have pointed towards the intensity of imitative and duplication work that underpins GitHub code production (Lopes et al., 2017; Mackenzie, 2018), reporting everything from a third to 70% of code on GitHub to be a duplicate, depending on the research technique. The high rates of duplication have also raised concerns about using GitHub to study patterns of software production (Lopes et al., 2017).

In addition to being framed as problems for research, the imitative processes underpinning GitHub's mode of code production can also be turned into objects of investigation. Indeed, turning forking into a medium-specific method for studying imitative practices on GitHub, allows us to examine one aspect of platform specific forms of journalism coding.[57]

An analysis of forked and original repositories in my journalism code corpus, shows journalism initiatives to be participating in the platform's code ecosystem primarily with original work. Close to 80% of the over 3,500 repositories are created by these organisations and only 23% of them are copied or forked from other GitHub users and organisation, a lower rate than the numbers reported at the level of the entire platform, albeit through other methods (Lopes et al., 2017; Mackenzie, 2018). The lower rate that describes my corpus might be associated with the practices of organisation accounts, as opposed to those of individual user accounts. An examination of the types of users that journalism code accounts "fork" repositories from is also insightful as it enables an exploration of the fields that journalism code blends with and draws on. While journalism organisations "fork" code from other journalism initiatives and journalist-coders, imitative flows extend outside the domain of journalism and see news initiatives draw on code produced by developers from other domains as well as by online platforms such as Netflix, Airbnb and WordPress. The latter may be taken as another approach to explore how journalism is platformised through the integration of code produced by online platforms in news operations.

4.3.5 How Journalism Code is Engaged With, Valued and Ranked

Having illustrated one way in which imitative practices can be studied on GitHub and found that journalism initiatives in my corpus primarily engage in original work when seen from the perspective of the "fork" function, next I examine participation in journalism code. As discussed in detail in section 4.2, participation in GitHub is a socio-technical construction aligned with the

57 For other methods to study imitation on GitHub see, e.g., Lopes et al. (2017).

platform's economic imperatives, involving processes of tracing, counting, calculating, recommending, intensifying, multiplying, archiving and mining engagement in the production and evaluation of platform content, through multiple interfaces, back-end and front-ends.

For these reasons analysing participation in journalism code becomes a way of examining the forms of work that journalism repositories attract, as well as how journalism code is evaluated, ordered, ranked, recommended and organised on GitHub, based on various social counters and composite calculations. Studying which code is successful in attracting platform work and in rising to the top is important in order to understand the hierarchies of visibility and value that social counters organise. It can also help us understand the modes of valuation that these rankings serve, i.e. what journalism code is of value in the GitHub platform ecosystem and is recommended to others.

To explore the forms of work that the journalism code space attracts, I analyse the types of acts associated with original journalism code repositories, i.e. those that do not originate in a fork. GitHub's Events API end point makes available 37 event types that make up the platform's public timeline data, some associated with repositories and some associated with other platform objects. These are used to compose the various activity feeds and streams of the platform. The repository end point makes available a selection of these, in the list or count form: "stargazers count", "subscribers count", "forks count" and list of contributors and their "commits count" for a given repository. I briefly describe each platform specific form of coding work before discussing how it plays out in my collection of journalism repositories.

Users revise or make changes to their GitHub repositories (e.g. adding, deleting or modifying code or any of the files of a repository) through commits. Each commit is associated with a repository contributor. The repository end point of the API allows the extraction of the top 500 repository contributors based on their number of commits. In the case of the journalism code repositories examined here, none of the repositories reached that limit, as the highest number of contributors recorded for a repository was 132. In this

case the number of commits for a repository can be determined by totalling the commit counts for all repository contributors. As discussed in the previous section, users can also copy repositories by forking them in order to propose revisions to them or to start their own project according to open source principles (GitHub, n.d.-f). At the repository end point of the API, a forks count is recorded for each repository. Users can also engage in acts of "sociality". In the case of GitHub, being "social" towards a repository can involve monitoring changes to a project through the "watch" function (GitHub, GitHub, n.d.-e). At the repository end point of the API a "subscribers count" is recorded for each repository. Finally, users can "star" a repository to bookmark it and generate recommendations, or as an act of public appreciation (GitHub, n.d.-e). At the repository end point of the API a "stargazers count" is recorded for each repository.

Both the user actions and their counting and release via the API participate in the ordering of journalism code, i.e. in the organisation of degrees of visibility and how the content is made available to users through various recursive activity streams and feeds.

First, a comparative analysis of the volume of different kinds of social acts associated with journalism code repositories can provide an indication of the forms of work that journalism code repositories attract as well as of how journalism initiatives inhabit GitHub.

	REPOSITORY ACT	SUM OF REPOSITORY ACT
1	Committing	355,752
2	Watching	75,115
3	Starring	46,128
4	Forking	12,193

Figure 19: Repository act counts for 3,665 journalism code repositories until June 2016.

The dominance of acts of committing or revising repository files shows GitHub to be a site of production as repositories are not just posted and

frozen in the form of original posting but updated and modified. But GitHub is also a site of sociality or, rather, connectivity. Sociality on GitHub is instrumentalised as public evaluation of performance, monitoring, circulation and imitation, as journalism code repositories also register acts of starring, watching and forking. These may also be associated with what in the context of journalism has been discussed as participatory transparency (Karlsson, 2010), namely the opportunities for the user to become involved in the news process through the possibilities for user action afforded by the technical architecture of digital platforms.

Furthermore, there is a great variation within this space, albeit following a common pattern. Across all repository acts, be they production or sociality, we encounter a power law distribution whereby a small number of repositories attract the largest quantity of platform work, contributors and attention, and a large number of repositories attract very little to no work and attention.

Almost half of the repositories have a single contributor thus attesting to individual modes of work active in the journalism code space but perhaps also to other kinds of labour required to attract contributors. Indeed, while numerous studies have been dedicated to studying various aspects of collaborative software development on GitHub, individual modes of public coding or single authored repositories might also be worth investigating, if only because they might make up a large part of the platform's code ecosystem, as this study of journalism code repositories seems to indicate. Close to half of the repositories do not receive any stars and are thus public but practically invisible as the absence of events leaves them out of the recursive recommendation streams in which high event repositories are included.

This power law distribution is common for social media activities (Mackenzie et al., 2015; Marres, 2017a). Power law distributions are produced by device cultures which deploy user activity to organise visibility, order content and recursively deliver recommendations back to the user in the form of various kinds of rankings. Many of these actions and social counters participate in a number of platform ordering and valuation mechanisms, such as "trending" repositories and the ranking of developers. GitHub platform developers

describe the composition of the trending feature for repositories and developers as follows: "We look at a variety of data points including stars, forks, commits, follows, and pageviews, weighting them appropriately" (Rohan, 2013). This means that not all repositories receive the same degree of visibility on the platform and repositories that receive more stars will be given more visibility through generic platform streams and personalised feeds.

	REPOSITORY	STARGAZERS COUNT
1	nytimes/objective-c-style-guide	4,050
2	BBC-News/Imager.js	3,689
3	fivethirtyeight/data	3,648
4	BBC-News/wraith	3,556
5	guardian/scribe	3,406
6	guardian/frontend	3,208
7	newsapps/beeswithmachineguns	3,206
8	Quartz/bad-data-guide	2,477
9	nytimes/pourover	2,402
10	Quartz/Chartbuilder	1,813
11	freedomofpress/securedrop	1,713
12	documentcloud/visualsearch	1,664
13	nytimes/backbone.stickit	1,655
14	nytimes/NYTPhotoViewer	1,529
15	propublica/upton	1,491
16	nytimes/ice	1,472
17	nytimes/gizmo	1,416

Figure 20: List of top seventeen journalism code repositories according to their stargazers count in June 2016, that together receive more than half of the total number of stars in the collection.

Indeed, more than half of the stars go to the top seventeen repositories (see Figure 20). But this ordering of repositories is not just the outcome of

platform technicity[58] and its modes of organising visibility. The ordering produced by the star counter shows that even in the context of domain specific inhabiting of the platform, what is valued are repositories that conform to the platform's primary asset form, source code, and that address themselves to one of the platform's primary stakeholder, developers. Moreover, a feature of these repositories with a high star count is a concern with materials that sustain developers and their software development work, as opposed to open source software end users as was the case with SourceForge, the other online coding platform. This reminds of Kelty's (2008) notion of free software developers as recursive publics, relevant in this context because it draws attention to the fact that what is valuable is the production of materials that enable programming work to be sustained and performed. Illustrative of this is the repository which received the highest star count at the time of data extraction, a style guide developed by *New York Times* developers containing coding styles and conventions for the Objective-C programming language. Others include testing tools to enable developers to catch bugs in their programmes, server load testing utilities, software libraries and utilities to build linkages and enhance software libraries. Alongside valuing of content aligned with the platform's primary asset form and end-user group, domain specific valuation registers are also visible in highly starred content that addresses itself to journalists and non-programming publics. This would include a collection of datasets that underpins reportage at *FiveThirtyEight*, a guide for dealing with common problems in data reportage, and the source code for a whistleblower platform. This attests to the extension of GitHub's model of social action (Mackenzie et al., 2015) to include other journalistic practices beyond coding. Producer tactics also play a role in content valuation as the top repositories are well documented, maintained and licensed.

4.3.6 Journalism Code as Ephemeral Construction

[58] While the concept of technicity has a longer history (see, e.g., the work of Martin Heidegger), in the context of digital social and media research technicity refers to the socio-technical relations or the articulations of software and users, human and non-human agents that characterise digital devices (Bucher, 2012a; Niederer & van Dijck, 2010; Weltevrede, 2016).

Finally, another aspect of platform-specific journalism coding that can be explored are the temporal dynamics of journalism code.

Freshness is a temporal dimension that is particularly valued in social media platforms. It is one of the organising principles of social media feeds, notification and activity streams and trending algorithms (Berry, 2011; Weltevrede, 2016). GitHub provides its users a news feed or activity feed organised by freshness where they can follow activities of the people and repositories they are connected with. The GitHub trending algorithm for repositories and developers also takes into account recency of events (Rohan, 2013).

Social media and search engine content date stamps present a number of social and media research affordances. Social and media researchers have repurposed date stamps to examine the pace of search engines and social media platforms, i.e. the temporal rhythm at which fresh content is introduced in their streams (Weltevrede et al, 2014). They have also used date stamps to examine variations or fluctuations of issue compositions (actors, topics, sources) over time, known as liveliness (Marres & Weltevrede, 2013). They have also used them to examine the effectiveness of censorship by taking the "freshness" or "staleness" of websites as an indicator. Here freshness or staleness is determined based on how frequently and recently sites have been updated (Weltevrede, 2016).

Similarly, repository date stamps can be repurposed to explore freshness as a distinctive concern of GitHub, as well as temporal dynamics of journalism coding, such as durability or ephemerality of repositories.

I will start with freshness. The first issue is to determine what constitutes a fresh repository. Measures of freshness are embedded in various platform features. For example, GitHub takes recency of events received by a repository as one of the data points in the calculation of trending repositories. While how exactly recency figures in the trending repositories calculation is not revealed by the platform, trending repositories are calculated for the last day, the last

week and the last month, which can be taken as an indication that freshness can be construed as receiving an event at the latest in the past month.

Keeping this platform specific qualification of freshness in mind, an examination of the dates when journalism code repositories had most recently been updated at the date of data extraction, indicates them to largely not be keeping up with the update culture built into the platform's recommendation features, to have distinct dynamics of update, or to be rather stale. A bit over 10% of journalism code repositories had been updated in the two months before the date of data collection.

Another temporal dynamic that allows us to move away from the preoccupation with "nowness" (Weltevrede, 2016), or what is happening now, of social media, is ephemerality or durability of issues and practices. In the context of GitHub, Mackenzie (2018) examines durability of repositories based on the number of events that they receive. Based on this measure, he describes ephemeral repositories as repositories that "flash into existence in the event of their naming before falling back into uneventful obscurity" (p. 45). Another way to examine the ephemerality or durability of repositories is by examining their date stamps. The repository section of the GitHub API returns a number of different date stamps for repositories, including the creation date and the date of the most recent update, or the last change made to the repository, recorded as a repository event. Calculating the time between these two dates indicates a significant number of repositories in this space to be largely ephemeral or short lived. A third of the original journalism code repositories only have activity in the month of creation (and an insignificant number of them have been created in the month when the data was collected). This suggests that they use GitHub as a site of storage and publicity and for projects with short timespans, rather than code development and collaboration over extensive periods of time. More than half of them had a lifespan between one month and two years, and the most durable repositories (eight in total) had a lifespan of over seven years at the time when the data was extracted (see Figure 21). These repositories may be deserving of further investigation to understand what makes a repository durable.

repository ▼	description
datadesk/latimes-calculate	Some simple math we use to do journalism.
datadesk/latimes-mappingla-geopy	A fork of the geocoding library geopy. Returns accuracy scores. Allows viewport biasing.
documentcloud/closure-compiler	A Ruby Wrapper for the Google Closure Compiler
documentcloud/cloud-crowd	Parallel Processing for the Rest of Us
documentcloud/docsplit	Break Apart Documents into Images, Text, Pages and PDFs
documentcloud/jammit	Industrial Strength Asset Packaging for Rails
propublica/table-fu	A utility for spreadsheet-style handling of arrays (e.g. filtering, formatting, and sorting)
propublica/vid-skim	Transcripts and commentary for long boring videos on YouTube!

Figure 21: List of most durable journalism code repositories, with a lifespan of over seven years at the time of data collection.

4.4 Conclusion

In this chapter I focused on GitHub as a news device. More specifically, I examined how the platform matters or makes a difference to journalism coding from a material-economic perspective, and what modes of studying these practices it may afford.

In relation to how the platform participates in news work, I captured its contribution by proposing to understand journalism coding that is public on GitHub as connective coding. Connective coding captures the participation of journalism code repositories in the platform economy as assets that have the potential to be variously capitalised by the platform (e.g. through future capital investments), and by its ecosystem of third-parties (e.g. through integration of platform data and knowledge derived from it into various third-party services that build on top of GitHub). The enmeshing of social and professional practices that use platforms with platform economic logics is of course not

unique to GitHub. The tensions and implications that emerge from this entanglement for power relations between institutions, domains of life, issues of labour, agency, public expression and information politics have been extensively discussed in studies of social media with a political economic angle (see, e.g., Gillespie, 2010; Langlois & Elmer, 2013; Nieborg & Poell, 2018; van Dijck, 2013).

Following the material orientation of the news device perspective, the contribution of this chapter has been to blend such approaches with a sensitivity towards the specificities of digital devices and describe the concrete medium-specific configuration of technical infrastructure and economic imperatives through which coding work is structured on GitHub. This extension of platform critique to an important but perhaps less critically investigated social platform can also be seen as a contribution to software and platform studies.

From a news research perspective, this chapter makes a contribution both to the study of social media and the news, and to the study of the role of programming, open source software and computation in journalism.

In relation to the study of social media and the news, I aimed to complement perspectives on coding on GitHub that understand them as social coding or "social open source" (to emphasise networking, collaborative and participatory aspects), with a perspective that emphasises the re-centralisation of journalism work outside the newsroom, as a platform asset. In doing so I hope to have also contributed towards a broader orientation for the study of social media and the news that does not take social media simply as backgrounds for understanding news work. Instead, this orientation would take as an object of study the platformisation of news and associated processes to render it an economically valuable part of the platform ecosystem. This orientation emphasises that to better understand the implications of social media platforms for news we should complement the internalist focus on their implications for professional practice with a focus on how these practices and their products are situated in the wider platform ecosystem, which the notion

of connective coding gestures at.

In relation to the role of programming, open source software and computation in journalism, the contribution is to approach journalism coding on GitHub as a platform-specific form of work that is enacted through material affordances, features and culture of the platform. I did so by using a medium sensitive approach that repurposes platform features to examine how journalism initiatives participate in the GitHub platform and how the platform is itself configured in this situated domain of practice. By repurposing the platform's "fork" feature for example, I found that journalism initiatives participate in the platform's code ecosystem with original work at a greater rate than numbers reported by other studies for the entire platform, while also engaging in imitative work. By repurposing collaboration and engagement metrics, I showed GitHub to be a site of journalism code production and not just of storage and transparency. By repurposing repository date stamps, I showed that many journalism repositories are largely ephemeral or "uneventful" in the words of Mackenzie (2018) and thus not conforming with the platform's update culture. An examination of top starred journalism repositories reflects a mix of modes of valuation: platform specific through the high valuation of materials that sustain software development work rather than software as an end product, and domain specific through the high valuation of non-code repositories such as datasets or data editorial and analytical guidelines. This attests to how, by entering the platform ecosystem, journalism objects also open themselves to platform-specific modes of valuation.

Finally, a methodological contribution is made by extending the study of interactions between GitHub and journalism coding to also include how the platform can be configured for news research and the participation of the platform in research methods and knowledge production. I illustrated techniques by means of which researchers can use the platform's analytical capacities at the same time as trying to modify their configuration to align them with their own research questions and interests. While platform analytics might push the direction of study towards that which is popular and trending now, the researcher needs to creatively work to adapt these for the study of social

questions that emphasise variability of practice and temporal dynamics other than the platform's preoccupation with the now.

While this study was primarily illustrative of conceptual approaches and research methods that can be brought to bear on the study of coding practices in journalism, future research may extend these analyses to larger collections of code repositories, and examine the potential of other categories of platform data for news research. Moreover, while this study illustrated how the platformisation of news can be examined through the lens of GitHub from the point of view of platformised journalism coding, platformising news occurs in various ways through multiple platforms, all of which are deserving of investigation. In the next chapter, I will address another aspect of how news infrastructures are platformised, through trackers embedded in news websites that insert news operations in various platform ecosystems and economies.

5. Making Audience: Web Trackers as Audience Marketplace Devices in Professional and Junk News

In the previous chapter I engaged with GitHub as a connective coding device. I discussed how GitHub platformises journalism coding and renders it amenable to financial capitalisation through interplays between its technical infrastructure and economic imperatives. In this chapter I take on another digital device that renders another type of news work economically valuable, namely third-party web trackers.[59] In doing so I expand the range of actors that matter in relation to digital journalism to include three other important ones in the discussion: audiences, the online advertising and marketing industries, and other forms of online content production known as fake news.

The "fake news" scandal is linked to increasing concerns about the dawn of a "post-truth era" (Sismondo, 2017), characterised by the unsettling of established epistemic hierarchies, the blurring of boundaries between news and other less reputable forms of digital cultural production and an increasing suspicion of experts and expert knowledge on the side of the public.

This concern about disruptions of established knowledge hierarchies is well illustrated by reports that false stories about the 2016 US presidential elections, packaged as news, outperformed "real" news in the domain of Facebook engagement rates (Silverman, 2016). That is, dubious facts had circulated more successfully and reached more individuals on this social media platform than journalistic facts about the US elections. This, alongside several other scandals, have brought the platform's reputation into decline.

But this is not the first time that the internet's reputation is questioned due to its association with deceptive knowledge practices. Rogers (2002) notes that

[59] This chapter is based on research conducted as part of a collaborative multi-institutional project I co-led with Jonathan Gray, Tommaso Venturini and Michele Mauri in 2017-2018 with support from First Draft and in collaboration with several journalists, media organisations, public institutions and others. The first phase of the research that underpins this chapter was conducted during a data sprint at the University of Amsterdam in March 2017 together with Michele Invernizzi and Mischa Szpirt. An earlier version of this study was published as a chapter in the research report *A Field Guide to "Fake News" and Other Information Disorders* (Bounegru et al., 2018). Another version of this research will be published as part of a article in *New Media & Society* (Gray, Bounegru & Venturini, forthcoming). For the purposes of this chapter the analyses were redone, extended and rewritten.

from its early days the web has raised concerns as "a medium of dubious repute" (p.192), a "rumour mill" where "rogue" websites and "rumour merchants" engage in spreading questionable and toxic content. These early concerns faded at least in part with the advent of search engines such as Google, which organised web sources based link authority logics, although they raised other concerns about the information politics of such ordering mechanisms (Rogers, 2004).

Today concerns about the role of online platforms in spreading dubious content have returned to public attention, and this time have brought social media and their like economy (Gerlitz & Helmond, 2013) under intense scrutiny. As colleagues and myself argue elsewhere, social media platforms should be seen as a matter of concern not only because of their failure to appropriately respond to "fake news" but, perhaps more importantly, because of their uncanny success in enabling the circulation, virality, monetisation and consolidation of such forms of digital cultural production, a phenomenon which elsewhere we discuss as "infrastructural uncanny" (Gray, Bounegru & Venturini, forthcoming).

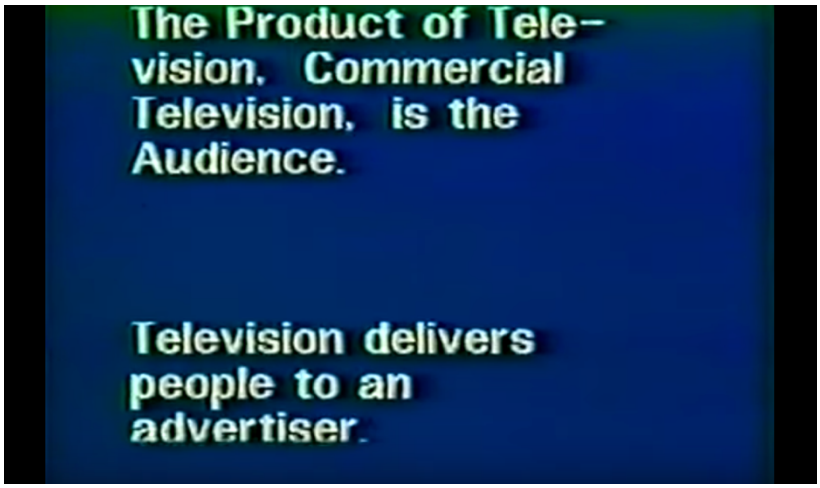
Indeed, the like economy, set up produce virality by intensifying the social validation of content and user engagement by means of social buttons and engagement counters spread across the web, is instrumental to such digital forms of "content farm[ing]" (Turow, 2011). Investigations exposed the profits that can be made from such forms of digital content production through an economic model based on the successful combination of low cost clickbaity sensationalist content which has the potential to go viral, with what in advertising is known as "click arbitrage": buying website traffic cheaply through advertising on platforms such as Facebook and bringing these audiences on the site, where ad revenue can be made through ad networks such as Google's AdSense (Nickel, n.d.; see also Dewey, 2016; Silverman & Alexander, 2016; Subramanian, 2017).

While debates about the economics of fake news have focused on exploring approaches to sanction disreputable publishers, such as by blacklisting them

from ad networks and flagging their content on Facebook, in this chapter I suggest that "fake news" can also be viewed as an "empirical occasion" (Marres, 2013; Marres & Moats, 2015). As an empirical occasion, fake news renders visible and amenable for social analysis not just the relations between *rogue* content producers, social media, audience participation and online advertising. It also renders analysable the relations between value creation, content production, social media, audience participation and online advertising that underpin digital cultural production *more generally,* including the more reputable kinds. In other words, "fake news" may be seen as an opportunity to explore the economics of various forms of digital cultural production.

Indeed, the dotcom crash has seen a shift in online business models towards advertising-supported models organised around the production and sale of audiences for most businesses with an online component (West, 2017). News publishers too have early on realised that people would not be willing to pay for news online and have turned to advertising models (Turow, 2011).

Richard Serra's iconic 1973 short film "Television Delivers People" reminds us that advertising-supported cultural production generates not just programming, entertainment and news but also the audience product. In the context of online news, the implementation of the EU General Data Protection Regulation (GDPR) in May 2018 has made visible again the interdependence between news production and audience tracking and commodification, as some news websites unable to guarantee lawful use of their audience's personal data have paused their services to EU citizens.



Los Angeles Times

Unfortunately, our website is currently unavailable in most European countries. We are engaged on the issue and committed to looking at options that support our full range of digital offerings to the EU market. We continue to identify technical compliance solutions that will provide all readers with our award-winning journalism.

Figure 22: (a) Screenshot from Richard Serra's iconic 1973 short film "Television Delivers People". (b) Screenshot from the LA Times' website after the enforcement of the GDPR, indicating the website is unavailable to users accessing the website from the EU.

While audience commodification has traditionally been the object of political economy and critical media research, in this chapter I link research into the economic underpinnings of news with socio-technical approaches to digital media, in order to examine the tracking infrastructures and practices that underpin audience commodification. I focus on two forms of advertising-supported digital cultural production, professional news and junk news. The question that drives this chapter is: *What can the news device approach contribute to the study of the audience marketplaces in which advertising-supported digital cultural production is embedded?*

I start by revisiting digital transformations of the media audience marketplace (Napoli, 2003, 2011). Next I introduce a news device approach to studying the

audience marketplaces in which different forms of digital cultural production are embedded, focused on the tracking infrastructures of websites. Finally, I discuss the results of an empirical visual network exploration of audience marketplace practices and tracking infrastructures across a set of professional and junk news sites, and their implications.

5.1 Digital Transformations of Audience Marketplace Practices

While audiences play multiple roles in news and their relations with journalism are multifaceted (for a discussion, see, e.g., Lewis & Westlund, 2015; Peters & Witschge, 2015), in this chapter the focus is on one particular aspect of this relationship, the audience as economic product or commodity. There is a long history in critical media and political economy research that examines the construction of the "audience commodity" as part of the analysis of advertising-supported media and cultural production (for an early account, see, Smythe, 1977). But in what follows I will primarily focus on the transformations to the audience marketplace brought about by the internet and digital platforms.

In the context of news, the turn from partisan newspapers to the penny press in the US in the 19th century marked a shift in media business models "from the sale of *products*—newspapers—to the sale of news corporations' *audiences* to subsidise media production" (West, 2017, p. 5). Today a great part of news production in many countries is advertising-supported (Nielsen, 2016). The audience product in the context of these advertising-supported operations is at its simplest understood as a "representation of consumer attention to advertising messages" (Napoli, 2003, p. 22). The production and exchange of this product is sustained by a marketplace arrangement which requires multiple actors and mechanisms to measure audiences, calculate and negotiate their value, rate publications based on their audiences and exchange audience attention for money.

The internet and digital platforms and services are linked to significant developments in the audience marketplaces associated with various forms of cultural production. According to Napoli (2011), the internet today represents a laboratory for the media audience marketplace. In this laboratory audience marketplace participants experiment with various programmatic instruments, techniques and methods to construct and deliver the audience product by tracing, measuring and calculating user characteristics, behaviour, tastes and preferences and anticipating future behaviour. They also test techniques to target and serve ads and measure their effectiveness, to assign value to both audience segments and websites, and, on the user side, to resist such measurements through anti-tracking and ad-blocking software.

The currency around which the audience marketplace is traditionally organised is the exposure metric, also known as "eyeballs" (Napoli, 2011). One of the first cases of online advertising is reported to be dating from 1994 when HotWired displayed an AT&T banner ad on its page (Evans, 2008; Turow, 2011). Early online advertising was organised around banners or display ads which were sold based on measures similar to those operating in traditional media such as the impression-based model of "cost-per-mille", or cost per one thousand individuals who were served the ad (Evans, 2008; Ratliff & Rubinfeld, 2010).

In the second half of the 1990s, a competing currency is established in the context of search engines, the cost-per-click (Evans, 2008). Associated with the rise of search engines and search-based advertising, this coin of exchange sees payments made only when viewers click on ads. The rise of search advertising has seen the revenue share from online display ads, used by the news media, declined by more than half in the first decade of the 21st century (Evans, 2008). The development of performance-based models such as the cost-per-click, and later of the cost-per-action, are part of what Napoli (2011) calls the "post-exposure audience marketplace", where, alongside exposure, other currencies, measurement instruments and audience conceptions are being experimented with and are shaping what the audience product is becoming.

The configuration and actor composition of the post-exposure audience marketplace is complex. While traditionally participants in this marketplace consisted of media publishers, advertisers (and related service providers such as advertising agencies), audience measurement companies and audience members (Napoli, 2003), today we are faced with an increasingly complex configuration of "interlocking multisided platforms" (Evans, 2008; for a discussion of multi-sided markets, see Chapter 4).

While traditionally news media were two-sided markets which provide readers with information and entertainment and advertisers with an audience for their product (Nielsen, 2018; Rochet & Tirole, 2003; van Couvering, 2017), today they are undergoing a process of platformisation, whereby they are joining other types of cultural producers in the precarious position of content providers in online platform markets, such as that organised by Facebook (van Couvering, 2017; Nieborg & Poell, 2018). In doing so they are becoming increasingly reliant on platform instruments to make audiences measurable, calculable, analysable and economically valuable.

At the same time, online advertising and marketing industries are becoming increasingly complex and are too becoming dependent on social media platforms such as Facebook (see, e.g., Helmond, Nieborg, & van der Vlist, 2017).

According to Nieborg (2016), in recent decades "the arrival of new intermediaries, the changing role of incumbents, and the adoption of internet-enabled mobile devices resulted in an increasingly opaque multi-sided market structure" (p. 4). In this complex arrangement exchanges and connections between the supply and demand side of advertising inventory are mediated by an increasingly large number of intermediaries and data brokers, both more established (e.g. audience measurement firm Nielsen) and more recent (e.g. online platforms and their marketing and advertising services). An important development is the emergence of advertising networks such as Google AdSense to monetise the "long tail" of the internet, i.e. small websites and user generated content (desilva + phillips, LLC, 2008; Gehl, 2014). Ad networks

rely on third-party tracking techniques to enable advertisers to reach audience profiles across a large number of sites, as well as programmatic tools for publishers to sell their media inventory, and for advertisers to buy it and place ads. A more recent intermediary, ad exchanges, centralises the selling and buying of ad space and takes its automation a step further through real-time bidding, the buying and selling of impressions, and the possibility to reach the user loading a website in real-time by placing bids in automated auctions. Indeed, any unreserved ad space on a publisher's website may enter the automated real-time bidding system of ad exchange service where ad networks, demand-side platforms or other ad exchanges can place bids (amounts of money they are willing to pay) to fill the ad space in real-time, as a user's browser loads a webpage (Interactive Advertising Bureau, 2014).

In this context, yield optimisation companies or supply-side platforms, such as Pubmatic or Rubicon, take on the role of supporting publishers in their programmatic interactions with advertising platforms by e.g. evaluating and filtering bids according to the publisher's criteria (Turow, 2011). On the advertisers' side, intermediaries include demand-side platforms, which provide media buying services to advertisers and mediate interactions with ad exchanges in the benefit of the advertiser.

Increasingly important in this ecosystem are also data intermediaries or data brokers such as BlueKai or Lotame. These are services that aggregate user information from multiple online and offline sources and offer or resell it to a number of other companies (Zuiderveen Borgesius, 2014), such as supply-side platforms. These use it to enrich user profiles made available to advertisers.

This complex market configuration is enabled by the stabilisation of cookie-based and other online user tracking mechanisms. Cookies are "small text files that sites place on a user's computer to identify the user's browser, computer operating system, IP address, and (if the user provides it) personal information" (Gehl, 2014, p. 105). They were originally developed to enable e-commerce sites to remember and record multiple items selected for purchase by a user in a virtual shopping cart (West, 2017). In the audience marketplace,

cookies enable publishers to keep track of visitors to their websites and enable ad networks to track users' browsing behaviour across website. Today it is common practice for websites to be paid for placement of third-party cookies on their websites to enable users to be tracked across the web (West, 2017). While the cookie is perhaps the most well-known web tracking device, web tracking can be variously implemented, including through web beacons or bugs, invisible one-by-one pixel graphic images that enable sites to transfer user data to third-parties such as ad networks, and the more persistent "flash cookies" (Gehl, 2014; West, 2017).

Such web tracking mechanisms have increasingly supported an orientation in the post-exposure audience marketplace towards behavioural targeting, highly controversial due to its implications for individual privacy online (see, e.g., Brotherton, 2012; Zuiderveen Borgesius, 2014). This advertising technique involves the "delivery of targeted advertising to different members of the audience based on their demonstrated patterns of media consumption or behavioral responses such as information requests and other possible advertisement responses, such as click-throughs or product purchases" (Napoli, 2011, p. 111). This approach shifts ad targeting from a focus on publications, to a focus on those individual profiles whose characteristics are highly valued by advertisers (Bakir & McStay, 2018).

From the point of view of the media's audience conceptions, these developments have been understood as an increasing "rationalisation of audience understanding" (Anderson, 2011a, p. 553), whereby audience perceptions have become "increasingly scientific and increasingly data-driven, with more impressionistic or instinctive approaches to audience understanding increasingly falling by the wayside" (Napoli, 2011, p. 11). More recently a *Digital Journalism* special issue captures this shift in perceptions with the notion of "measurable journalism", described as a "term that encapsulates the cultural and material shift to digital platforms capable of providing real-time, individualizable, quantitative data about audience consumption practices" (Carlson, 2018, p. 409).

The post-exposure audience marketplace where behavioural targeting techniques play a central role marks a shift not just towards more scientific conceptions of audiences, which have underpinned this market from the beginning, but towards *particular scientific visions* inscribed in audience measurement systems active in this marketplace. These rearticulate audience representation from "statistical abstractions" (Napoli, 2003) or portraits of audiences based on broad descriptive categories of shared characteristics, towards a form of "radical behavioralism that calculates society without representing it" (Cardon, 2016, p. 104). The latter focuses on traces that differentiate individuals not for the purposes of representation but in order to act on and influence their behaviour.

In this chapter I argue that the web tracking infrastructures that materialise these rearticulations of audience understanding can be repurposed to explore in more detail the post-exposure audience marketplaces in which different forms of digital cultural production are embedded.

5.2 A News Device Approach to Audience Marketplaces

5.2.1 Web Trackers as Digital Objects

As discussed briefly in the previous section, in the context of advertising-supported digital cultural production, digital objects such as cookies and other web tracking devices play an increasingly prominent role in audience measurement and other practices that make up the audience marketplace. As snippets of third-party code to be found in the source code of websites, trackers form an invisible "data mining infrastructure" whose role is to establish connections between websites and various third-party services, and to enable data flows between them (Gerlitz & Helmond, 2013). Mayer & Mitchell (2012) see web tracking as part of "the increasing trend of third-party websites recording and analyzing users' browsing activities across unrelated first-party websites" (p. 1). Roesner, Kohno, & Wetherall (2012) give the example of "a website (like doubleclick.net) that has its tracking code included or embedded

in another site (like cnn.com)" (p. 1). Embedded in websites such as cnn.com, data flows enabled by such trackers may include recording visits to the website, extracting various kinds of user data, combining it with data from other databases, serving ads and measuring ad effectiveness. Trackers are not just invisible devices but they are also *dynamic* or *lively*, in the sense that user activities shape the data flows which are being initiated and their intensity (Gerlitz & Helmond, 2013). From this point of view, they should be understood not just as snippets of code in the source code of websites or as files on users' computers in the case of cookies, but as mechanisms that enable communication between the website, the user and third parties (van der Velden, 2018).

The attachment of this *invisible and dynamic "data mining infrastructure"* (Gerlitz & Helmond, 2013) to a website is linked to the changing nature of the websites, from being "designed and hosted by a single person, group, or organization" to being "increasingly composed of content from myriad unrelated 'third-party' websites in the business of advertising, analytics, social networking, and more" (Mayer & Mitchell, 2012, p. 1). Similarly, Helmond (2017) captures this transformation of websites in the social media age from "self-contained units" to "assembled units":

> The website can be seen as an assemblage of modular elements that on the one hand enable interactions with other actors on the web and on the other hand permeate or redraw the boundaries of the website by setting up data channels for the exchange of content and data stored in external databases. (p. 6)

These controversial web tracking devices and the associated data mining practices that they enable are extensively studied and problematised in relation to a number of issues: online surveillance, privacy and security concerns, and related policy challenges (Binns et al., 2018; Englehardt & Narayanan, 2016; Libert, 2015; Mayer & Mitchell, 2012; Soltani, Canty, Mayo, Thomas, & Hoofnagle, 2009; van der Velden, 2014), web economies, expanding data industries, data capitalism and associated issues of uncompensated digital labour and power asymmetries (Andrejevic, 2014; Gerlitz & Helmond, 2013;

West, 2017), web history (Helmond, 2017), forms of discrimination and bias that may emerge from the processing of such traces (Barocas & Selbst, 2014), comparisons of tracking practices across domains and countries (Castelluccia, Grumbach, & Olejnik, 2013; Deville & van der Velden, 2016; Libert & Nielsen, 2018), as well as practices of resistance in these tracking environments (Brunton & Nissenbaum, 2013). This diversity of issues attached to trackers and tracking practices reflects their highly contested and "multi-valent" (Marres, 2011) nature, in the sense that a number of different spheres are co-articulated through them, from audience economics, to innovation in ad and information delivery (e.g. personalisation of both news and ads, recommendations), editorial decision-making and consumer surveillance. In this chapter however I focus primarily on web tracking as a device for understanding audience marketplaces.

5.2.2 Trackers as Audience Marketplace Devices

Many (although not all) of the trackers part of the invisible data mining infrastructure of advertising-supported websites are associated with the actors and practices of the audience marketplace. This fact draws attention to how the participation of news publishers in the audience marketplace is predicated on and shapes the very material infrastructure of news, the website. This changing character of websites, combined with infrastructural approaches in new media and internet studies (Helmond, 2015b; Plantin et al., 2016), may open up new research possibilities in the area of news audience economics (Napoli, 2003). These would see the news websites be treated as an object not only for the study of various aspects of digital news production, distribution and recommendation (see, e.g., Boczkowski, 2004a; Bødker & Brügger, 2018; Matheson, 2004; Karlsson, 2010; Stroud, Scacco, & Curry, 2016), but also for the study of particular aspects of the making of audiences into economically valuable products, such as the actor composition of audience marketplaces of a website and its evolution over time, comparative studies of audience marketplace configurations across different types of advertising-supported cultural production, and the relationships between news institutions and other

industries active in the audience marketplace. Audience marketplace configurations in this context can be understood as assemblages of actors that participate in audience commodification through tracking devices that facilitate tracking, calculating, analysing, evaluating and monetising various kinds of user data and online activities.

Drawing on Callon & Muniesa (2005), I conceptualise web trackers as online audience marketplace devices to draw attention to the work they do to materialise relations, exchanges and data flows between different participants in the online audience marketplace.[60] Participants include users, digital cultural producers, advertisers and advertising-related services, audience measurement companies and other data intermediaries. While audience marketplace devices may include traditional forms of measurement that have been migrated to digital environments such as surveys and focus groups, in this chapter I make a contribution towards understanding "natively digital" audience marketplace devices, i.e. those that are specific to the web medium as opposed to those that emulate more established methods of measurement (for more on this distinction see, Rogers, 2014). Trackers can be seen as audience marketplace devices not just in the sense of mechanisms that facilitate transactions and exchanges between different participants in the audience marketplace (e.g. capture of user data, delivery of advertisement, delivery of attention to advertiser, etc.), but also in the sense of acting as "algorithmic calculative devices", which make various forms of algorithmic calculation possible (Amoore & Piotukh, 2016, p. 18; see also Callon & Muniesa, 2005). They do so by making user activities calculable by transforming them (e.g. the user gaze) into "locatable objects" (Introna, 2016) which can be detached, put in relation to other objects, calculated and assigned value (e.g. as impressions and currencies), and thus made economically valuable.

According to Callon & Muniesa (2005), to capture different aspects of economic calculations and their consequences, the researcher can take as a starting point the calculative agents who are active in the marketplace, the

[60] I am grateful to Tommaso Venturini for pointing me towards the work of Michel Callon and Fabian Muniesa on markets as calculative devices in this context.

construction of goods (in the case of my chapter, the institutionalised audience), or the calculated exchanges (in my case, the exchanges consist of the delivery of user attention and data to the content provider and to the advertiser), and I would add, controversies around currencies or coins of exchange in this marketplace (e.g. cost-per-click and cost-per-action).

In this chapter I take the first aspect, pertaining to calculative agents, as a starting point. More specifically I take the tracking practices of two forms of advertising supported digital cultural production, news and junk news sites, as a way to examine the audience marketplace configurations, third-party actor compositions and relations specific to various forms of digital cultural production. In focusing on the composition of the online audience marketplace, I also follow political economy researchers who have drawn attention to the fact that media critique focused on aspects of political economy should complement interrogation of scientific techniques and methods for audience construction (usually studied around the most well-known audience measurement firms, such as Nielsen), with studies that pay attention not just to a few leading actors but to *"the many systems available in the market*" (emphasis mine), i.e. studies of industries active in audience construction, and the relations between actors in these industries (Meehan, 1984, p. 218).

An approach that does not reduce the implications of tracking infrastructures to consequences for journalism practice seems particularly important in the case of news website tracking, because these sites have repeatedly been found to be one of the website categories with the largest volumes of online tracking activity (Englehardt & Narayanan, 2016; Libert & Nielsen, 2018). In a recent study focused on news sites in several European countries, Libert and Nielsen (2018) found that the volume of third-party tracking domains was consistently higher for news sites than for any other popular sites across the countries in the study. These tracking practices have significant implications outside journalism as well, as tracking activities facilitated by these sites feed the growing and controversial data industry (see, e.g., West, 2017).

Finally, the focus on arrangements of tracking agents is also important in addressing political questions pertaining to who tracks and analyses, and the implications of these activities. According to Callon and Muniesa (2005) calculative processes "are all costly activities that raise the question of calculative power" and of "asymmetries of calculation", linked to the uneven distribution of calculative equipment (p. 1232). Similarly, in the case of web tracking and related data mining processes, Andrejevik (2014) has drawn attention to the asymmetries of participation in online data mining between "sorters" and "sortees", i.e. "between those who collect, store, and mine large quantities of data, and those whom data collection targets" (p. 1673).

5.3 Studying Audience Marketplaces With Tracker Signatures

My method consists of tracing the relationships between forms of digital cultural production and audience marketplace actors by examining the third-party tracking mechanisms embedded in websites and visible in their source code (drawing on an approach documented in Helmond, 2017). While in the audience research industry tracking activities have begun to be used to rank and evaluate media company performance as an indicator of their capacities to capitalise their audiences (Napoli, 2011), in this chapter I aim to rework tracking activities away from such quantitative evaluative practices and towards a qualitative exploration of tracking practices. This includes paying attention to the relationships between changing website infrastructures, the technicity of trackers and their cultures of use across different forms of digital cultural production, as well as legal considerations with regard to third-party tracking.

For this analysis I take as a starting point a corpus nineteen English language mainstream news sites and nineteen junk news sites whose election stories were found to receive high Facebook engagement scores in the months before the 2016 US presidential elections (Silverman, 2016).

Following Helmond's (2017) suggestion that different page types might be

inscribed with different trackers, I include in my corpus the homepage as well as a selection of five popular article pages for each site, resulting in a corpus of 228 URLs. To identify the articles that receive the highest social media engagement scores I use the social analytics tool BuzzSumo.[61] I focus on the tracking practices of popular sites and pages because the interest (and concern) in "fake news" was prompted by the viral character of these stories, as well as because popular sites tend to be richer in tracking elements (Macbeth, 2017). While tracking-poor sites are also relevant to examine, for the purposes of this chapter I focus on forms of digital cultural production that are tracking-rich. The articles included in the corpus span mostly 2016 and 2017, with a few junk news articles dating back to 2015.

To explore the third-party tracking networks associated with these forms of digital cultural production, I use the Tracker Tracker (Digital Methods Initiative, n.d.-b), a research tool that identifies third-party data-tracking technologies in a given collection of websites. The URLs were run through the Tracker Tracker tool in March 2017.[62] Given the dynamic character of tracking, this discussion thus reflects third-party tracking domains loaded by the website at that time for a user who accesses the websites from Europe. Convenient as it might be for the researcher, the phenomena examined in this chapter are by far not stable objects of study. Junk news sites active around the 2016 US presidential elections are ephemeral constructions and due to the ensuing backlash many of them were retired not long after their stories became viral. Moreover, their tracking practices are also fluctuating, as responses to their viral character include blacklisting some of them from ad networks. Hence the audience marketplace configurations discussed in this chapter should be seen not only as economic-material arrangements but also as the outcome of public pressure and policy interventions. The instruments used to study them also shape the resulting picture of the phenomenon, as I will discuss next.

61 http://buzzsumo.com/

62 Preliminary work that informs this chapter has been conducted during two collaborative research projects which I led at the Digital Methods Initiative Winter School in January 2017 and at the Fake News Sprint in March 2017, both hosted at the University of Amsterdam.

5.3.1 The Tracker Tracker as a Research Device

Prompted by Brunton and Coleman's (2014) remark that a good research narrative includes not just a picture of the phenomenon as seen through the results of its study, but also the picture's "own blind spots, occlusions, and *range of focus*" (p. 79, emphasis mine), in what follows I will discuss a few aspects of my method that I think deserve further attention, with particular emphasis on the calibration mechanisms that make up the range of focus of this study.

As a media and social research tool, the Tracker Tracker repurposes the detection and classification capacities of the popular Ghostery privacy protection browser extension. In doing so, as discussed in Chapter 2, it is situated amongst device approaches to digital media research that seek to repurpose the analytical capabilities of existing digital services such as search engines and platforms (Rogers, 2013; Weltevrede, 2016). More specifically it is associated with a more recent phase of internet studies which aims to shift from studies of hyperlinks and single social media platforms to cross-platform analysis (Helmond, 2017; Rogers, 2018). In this case cross-platform analysis focuses on digital objects embedded in websites that create connections with third parties. Indeed, the Tracker Tracker tool has been used for a number of media research as well as journalistic projects (see, e.g., Deville & van der Velden, 2016; Gerlitz & Helmond, 2013; Helmond, 2017; Silverman, Singer-Vine, & Vo, 2017; van der Velden, 2014). Such approaches may also be thought of as "inventive methods", in the sense that they are characterised by a "variety and variability of purpose" (Lury & Wakeford, 2012, p. 5) and they may be put to use for a number of different research and other purposes.

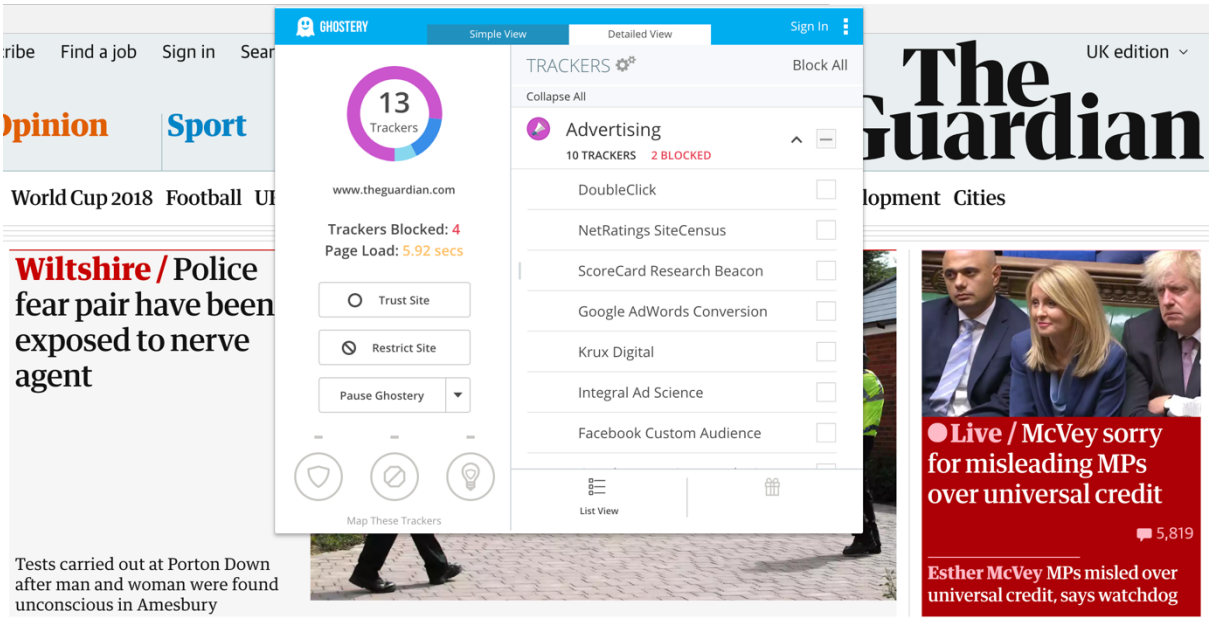


Figure 23: Screenshot of the Ghostery browser plugin visualisation of tracking elements active on The Guardian's homepage.

As a research device, the Tracker Tracker tool channels Ghostery's tracker detection and classification capabilities to enable researchers to extract third-party tracking elements for a single or a set of URLs in table or network format. While web tracking can be set up by first parties too (the website the user is voluntarily visiting), Ghostery focuses on third-party tracking elements embedded in first party websites as these are domains that the user connects to involuntarily.

Ghostery maintains a database of over two thousand trackers and over four thousand scripts associated with these trackers,[63] which it describes as one of the largest tracker databases.[64] Each tracker in the database has a URL and a profile.[65] The service matches requests to third parties sent during a page load against patterns, scripts or code snippets in the Ghostery database (e.g. "\.googlesyndication\.com\/simgad\/|\.googlesyndication\.com\/pagead\/|partner\.googleadservices\.com\/gampad\/" for Google AdSense). This allows the service to identify third-party tracking elements with which the page establishes connections (e.g. AdSense) and the companies that own them (e.g.

63 https://www.ghostery.com/submit-a-tracker/
64 https://ghostery.zendesk.com/hc/en-us/articles/115000734653-View-all-Trackers
65 E.g. see the tracker profile for the ad network Google AdSense: https://apps.ghostery.com/apps/google_adsense.

Google). A third-party element can have one or multiple scripts embedded in a web page (and thus can establish one or more connections to a web page), and a company can own one or more third-party elements. Such is the example of Google who owns multiple audience marketplace services that make use of tracking elements, including AdSense, Google Analytics and DoubleClick.

The Tracker Tracker tool enhances Ghostery by enabling researchers to compare the presence of trackers across websites (van der Velden, 2018) and to study "tracker networks" associated with websites (Helmond, 2017). The analysis of tracking networks builds on earlier forms of repurposing of natively digital objects to examine how web pages are associated, valued and ranked in online cultures, through e.g. hyperlink analyses, URL-hashtag analyses in the case of Twitter (analyses of associations between URLs based on hashtags), URL-page analysis in the case of Facebook (analyses of associations between URLs based on Facebook pages that share them). The Tracker Tracker also repurposes the tracker classification that Ghostery produces for the purposes of raising individual awareness of surveillant technologies and providing privacy protecting technologies and services.[66] Given that Ghostery's database, known as Ghostrank, is compiled based on tracker data collected from users who have opted to share information with the service (Macbeth, 2017), it means that Ghostery will be less effective at detecting less frequently used tracking elements (Englehardt & Narayanan, 2016).

The lively or dynamic character of tracking discussed in the previous section also shapes the study of tracking practices. Indeed, the data flows that are set in motion and captured for analysis depend on a number of things. This includes the location of the user who accesses the site. For example, the BBC does not serve ads for users visiting its site from the UK but does serve ads for users visiting it from outside the UK (BBC, n.d.). Other studies have noted that additional tracking may be activated upon user activities such as clicking on an ad or a social widget (Roesner et al., 2012). Advertising networks also contribute to the fluctuation of trackers loaded on a page, as they load ads (and

66 In the course of my doctoral research Ghostery was acquired by Cliqz, a German company producing anti-tracking technologies and services. See: https://www.ghostery.com/blog/ghostery-news/ghostery-acquired-cliqz/

associated tracking elements) from various suppliers who bid to display an ad based on the profile of the user who visits the website (Macbeth, 2017). This means that different page loads may result in different trackers being detected (Libert, 2018). Finally, certain types of third-party trackers such as ad networks act as hosts or aggregators for several other third-party tracking elements which will not be directly embedded in the first-party website but are requested or referred to by embedded trackers (Roesner et al., 2012). From the point of view of the publisher, such behaviour results in opacity, as publishers might not be aware of all third parties that access user data through their site (Joseph, 2018).

Finally, the configuration of the Tracker Tracker tool itself shapes the detection of tracking practices. Pages are loaded through a PhantomJS browser without a user interface, which means that any prompts for user input are ignored, which may impact the trackers loaded on the page. This would include prompts to accept cookie policies and to log into a website to be able to access its contents.[67] For this reason this technique may be seen to come closer to other lower bound methods likely to under-detect tracking activities (such as those used by Englehardt & Narayanan, 2016; Libert, 2015; Roesner et al., 2012).

5.4 A Visual Network Exploration of Audience Marketplace Configurations From the Perspective of Web Tracking

The network files generated by the Tracker Tracker tool for the lists of junk and mainstream news sites (homepages and articles) are uploaded and appended in Gephi,[68] a visual network exploration tool which is widely used in digital humanities, digital social research and other fields. This operation results into a network of 504 nodes and 5,897 connections. Of these, 219 are first-party URLs[69] (113 pertaining to junk news sites and 106 to mainstream news

67 https://tools.digitalmethods.net/beta/trackerTracker/
68 https://gephi.org/
69 The Tracker Tracker tool returned no output for a few URLs, particularly homepages of mainstream news sites, which reduced the original collection of first-party URLs to 219.

sites), and 285 are tracking elements, which makes the network a bipartite network. While relations between junk and mainstream news have previously been considered, among other things, in terms of their comparative valuation on social media in terms of engagement (Silverman, 2016), the comparative speed of their spread on Twitter (Vosoughi, Roy, & Aral, 2018), their comparative degree of reliability (Les Décodeurs, 2017) and their linking behaviours (Venturini et al., 2018), in this chapter I explore connections between them based on their invisible tracking and data mining infrastructures.

Before moving into visually exploring the network in more detail, a first observation can be made based on the volume of tracking elements and connections, which points towards the *distributed nature of audience commodification*. Indeed, while these practices have always been a collective accomplishment, this analysis reveals it as accomplishment distributed across an increasingly large number of inter-dependent actors. Overall there are close to 300 tracking elements in the network that receive close to 6,000 connections from first-party websites. Audience construction and monetisation is not enacted through standalone tracking activities but it is distributed across inter-dependent tracking elements, which often collect multiple user data points across multiple websites and communicate and exchange data with each other. Indeed, partnership programmes are a common feature in the online advertising industry (see, e.g., Helmond et al., 2017). For this reason, it is important to explore relations between actors in this market and to examine audience commodification not as the outcome of one audience measurement systems but as the outcome of interactions between multiple such systems. As mentioned in section 4.1, in this increasingly complex configuration in which exchanges and activities are increasingly managed programmatically, we are not only faced with an increasing "rationalisation" of ways in which the news media understands its audience (Napoli, 2011), but also with an intensification of audience measurement and analysis, and with particular forms of rationalisation that draw on "radical behaviouralism" (Cardon, 2016).

Visual network exploration (Venturini et al., 2015), is particularly suited for the qualitative exploratory analysis of tracking practices grounded in the empirical

world because it allows the researcher to examine individual actors and associations between them at a disaggregated level. While visual network exploration has seen a "renaissance" over the past decades, early key figures of social network analysis such as Moreno (1977) also pointed towards the importance of network visualisation as a method of exploration in social research (on the visual exploration of networks see also Venturini et al., 2017, 2018; on the methodological reflexivity required when using data visualisation as an analytical device, see Gray, Bounegru, Milan, & Ciuccarelli, 2016).

To visually explore the audience marketplaces in which different forms of advertising-supported digital cultural production are embedded, I use a "force-directed spatialization" technique which "simulates a physical system in order to spatialise a network. Nodes repulse each other like charged particles, while edges attract their nodes, like springs" (Jacomy, Venturini, Heymann, & Bastian, 2014, p. 2). This way, the disposition of the nodes in the space of this network representation gains meaning and can be interpreted as "a proxy of their structural similarity: two nodes being the closer the more directly or indirectly they are connected" (Venturini et al., p. 4). The outcome of this technique is the visual representation of relations between trackers and websites in the form of a network composed of regions of various node and edge densities (also known as clusters) separated by empty or sparsely populated areas known as structural holes (Burt, 1992). Clustering and the structural holes that separate them are illustrative of asymmetrical associations between actors in the network.

This technique is particularly useful for the visual exploration of tracking practices because its resulting clustering is similar to the outcomes of cluster detection or community detection techniques (Noack, 2009). Drawing on earlier work by colleagues and myself, I see visual network exploration as an iterative process involving "a constant toing and froing of categorization and observation, typology and topology" (Venturini et al., 2018, p. 269).

Given the (relatively) large number of nodes in the network for a qualitative analysis, to facilitate exploration I size nodes according to their number of

connections. Because the network is directed, i.e. links have directionality (in this case from third-party trackers pointing to websites), they can be sized by in- or out-degree, i.e. the number of connections they receive and the number of connections they send, respectively. Which measurement is relevant depends on the aims of the research. If the focus is on investigating individual news publishers and the connections that they establish with third parties, then nodes can be sized by in-degree to highlight the publishers which receive most connections from third-party trackers. In the case of my research, the question does not pertain to tracker ecologies at the level of the individual websites. For this type of analysis to yield interesting results, different website selection criteria than the ones used for this study might be better suited. Instead, my aim is to explore tracking practices across different forms of digital cultural production. Because of this interest, I size nodes by out-degree in order to be able to interpret the prominence of trackers across these information spheres.

These operations having been completed, the reader may notice from Figure 24 below that the network is not divided into two equally sized regions, which we could be expecting given that our starting points were two equal sets of URLs. Like many real-world networks, the layout of the audience marketplace network at first sight does not seem to present a distribution of nodes into any number of neatly separated clusters. Instead, the first characteristic we may notice is a large component at the bottom and a small one at the top which suggests an uneven distribution of associations between nodes, in need of further exploration.

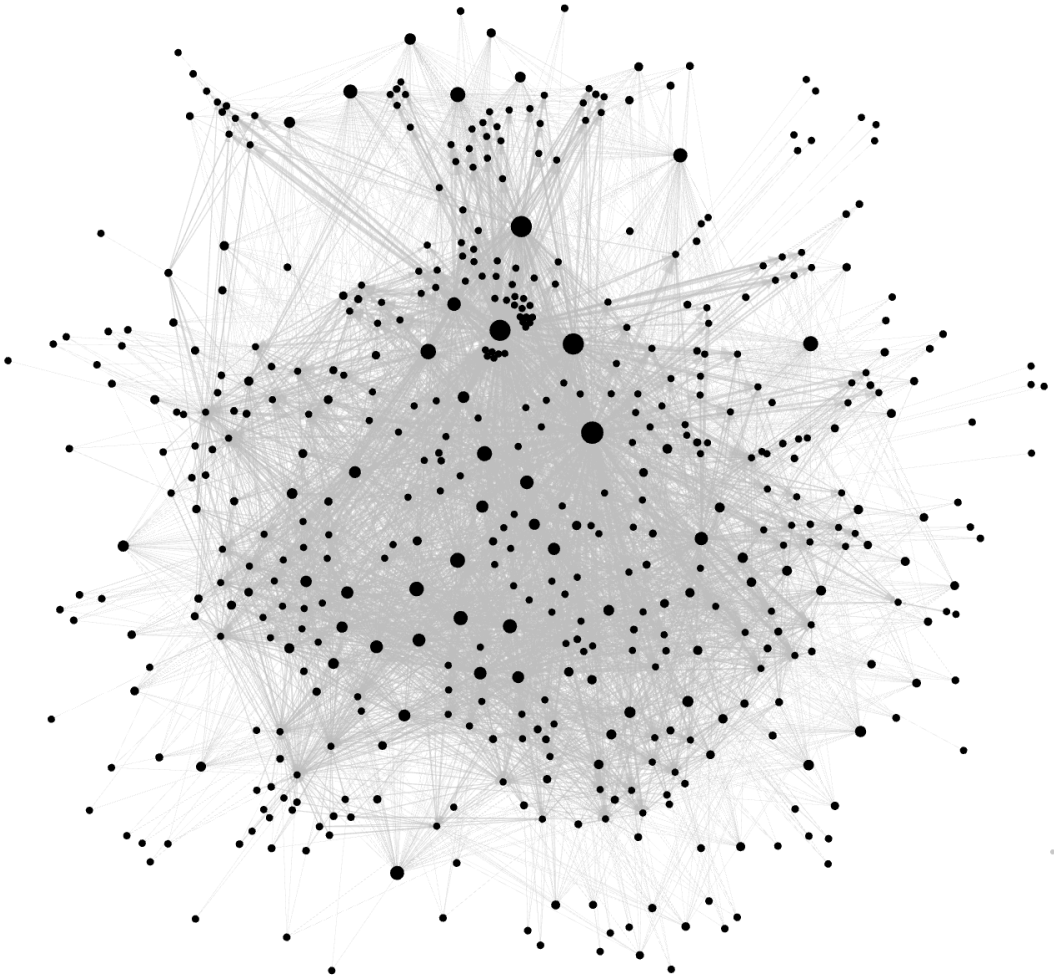


Figure 24: Mainstream and junk news sites and their third-party tracking networks spatialised with a force-directed layout algorithm in order to illustrate structural similarities visually, as node proximities.

In this case we could assume that this particular spatialisation of the network is indicative of differences in tracking practices associated with the two different types of URLs which were the starting points of this analysis: mainstream news and junk news. For the time being the original binary classification of URLs in the corpus will be used. As we will see, the classification will be revisited and enriched later in order to support the network exploration.

And indeed, if we colour nodes of the first-party domain type by the kind of websites they belong to and highlight outlinks sent by third-party domains by the type of their target first-party sites, we see that this website classification can explain the structure of the network, but it can do so only partially (see

Figure 25). While the majority of the URLs in the small component seem to be junk news URLs, the big component is divided between junk and mainstream news URLs.

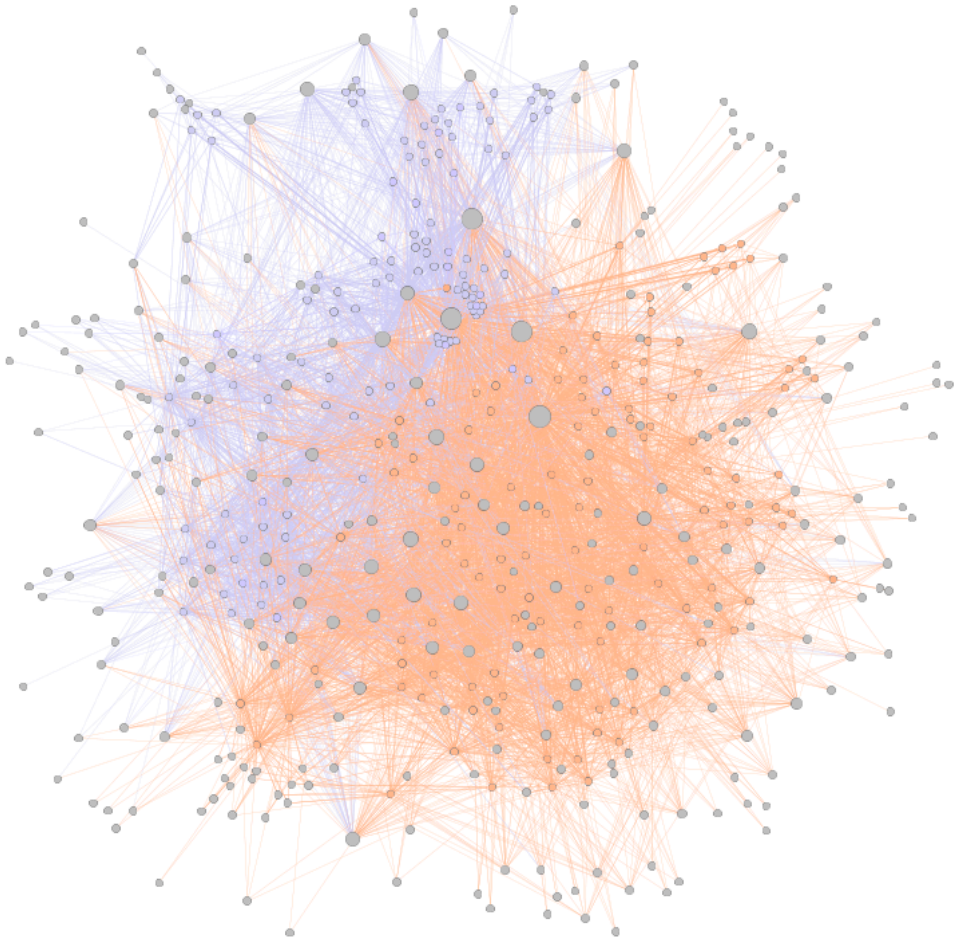


Figure 25: Mainstream and junk news sites and their third-party tracking networks, where junk news URLs are represented as purple nodes and mainstream news URLs are represented as orange nodes. Edges that link first-parties to third-parties are coloured by the first-party domain type.

Nevertheless, there are things we can learn from this analytical operation. Overall third-party trackers (of which there are a total of 285) receive more connections from mainstream news sites (3,763) than from junk news sites (2,134). News sites also connect to a larger number of unique tracking elements (243) than junk news sites (144). Mainstream news websites thus appear to be both more intrusive and more diverse in their tracking practices.

But, as I will return to later in this section, this may be taken as indicative of the frailty of business models and revenue streams of advertising-supported news media, increasingly pressured to monetise all their inventory.

Returning to the network exploration, as the URL classification does not fully explain the spatialisation of the graph, further analysis of the graph structure is needed. A closer look at the node composition and the types of connections making up the large graph component, shows that it consists of both mainstream news and junk news URLs, thus pointing towards a network composed of not two regions (around mainstream and junk news sites), but of multiple regions and thus tracking styles. When examining the densities of nodes in this component more closely, at least three more regions become apparent in the large pole of the network, two organised primarily around junk news URLs and at least one organised primarily around mainstream news URLs. Gephi's community detection algorithm, applied as a node colour layer, seems to confirm these regions (see Figure 27).

To strengthen this topological interpretation, I experiment with further node categorisation and analysis. While so far the topological interpretation relied on the classification of URLs, could the typology of tracking elements add something to the interpretation of graph regions? To explore this question, I first turn to the typology of trackers that the Tracker Tracker tool outputs based on Ghostery's classificatory work. As Ghostery is regularly updating tracker categories, I update the network with the most recent tracker classification, which is the most relevant for the purposes of this study because it focuses on the purpose of the tracker. Ghostery's aim to increase individual awareness of services which access their data when they browse the internet, becomes immediately apparent in the tracker categories. These categories collapse the multiple services and functions at work in the audience marketplace into generic labels such as advertising, aimed at a non-specialist public. To test the robustness of this classification I manually check the tracker category against service descriptions on their own websites as well as other online source and in a few instances reclassify the tracker to reflect its main purpose and categorise tracking elements that do not receive a label from

Ghostery. In what follows I will discuss a few observations based on this analysis.

5.4.1 Advertising-Dominated Tracking Networks

Layering the Ghostery classification over the third-party services in this network representation suggests digital cultural production to be deeply entangled with the online advertising and marketing industries (see Figure 26). The audience marketplaces in which these forms of digital cultural production are embedded are dominated by advertising industry-related tracking, which make up over two thirds of the third-party domains in the network and are dispersed over all regions of the graph. Audience measurement instruments such as site analytics and social media tracking elements are following advertising industry-related tracking in terms of usage. This observation points towards the reliance of these information spheres on a great number of advertising related services, which may in turn be an indication of the precarity of these information spheres which need to resort to multiple services to monetise their inventory.

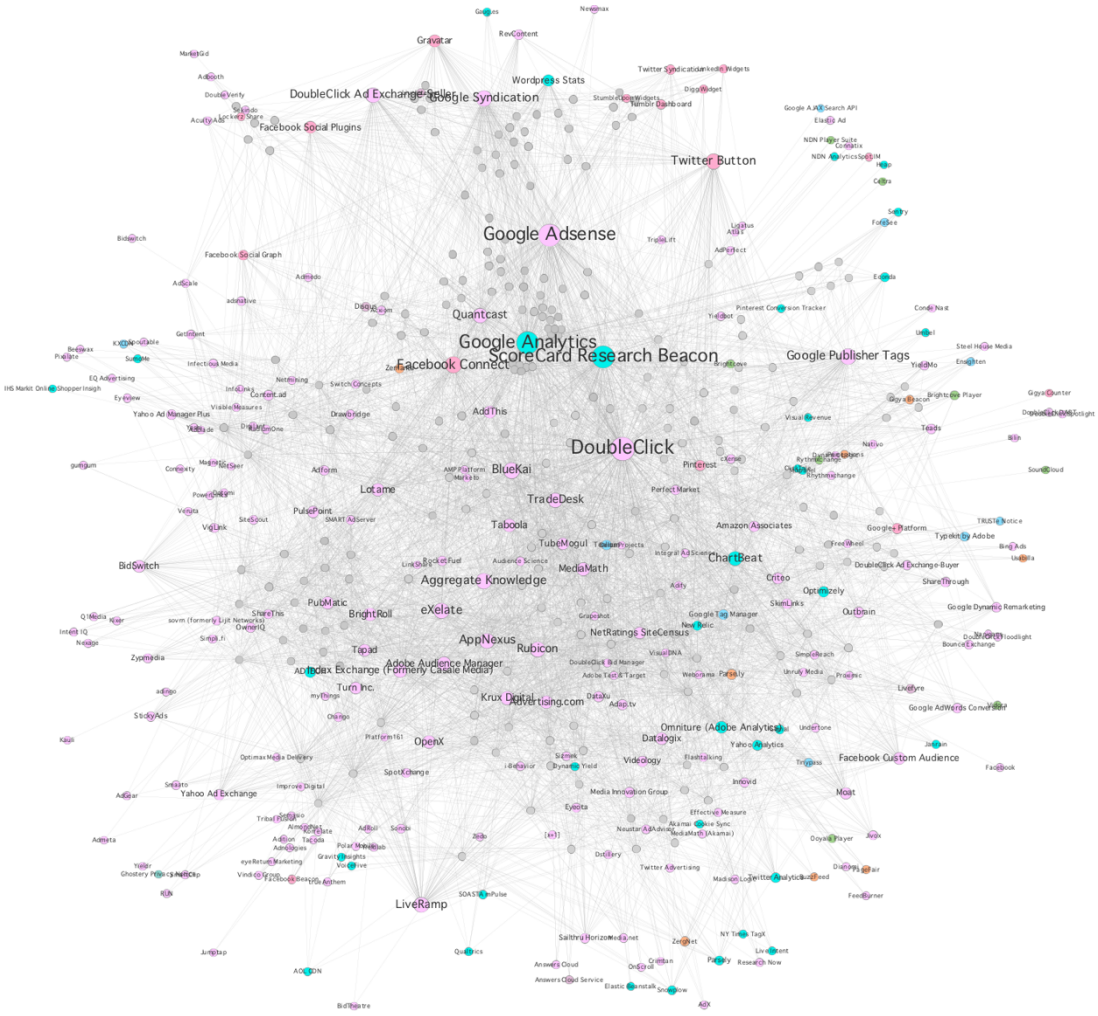


Figure 26: Mainstream and junk news sites and their third-party tracking networks, where third-party tracking elements are coloured by their type. Notable are the nodes in purple (representing advertising industry-related trackers), those in aqua representing in-site and cross-site analytics trackers and those in pink representing social media tracking elements.

Notable is also the bridging position of the in-site analytics service Google Analytics between junk news sites and mainstream news sites. This position indicates its centrality to both digital cultural production types. Indeed, Google Analytics has repeatedly been found to be one of the most widely used third-party domains across the web (Englehardt & Narayanan, 2016; Macbeth, 2017). But this external classification understandably does not correspond to the network structure and thus does not help to further elucidate the structure of the audience marketplace network, leaving space for further investigation.

To advance this analysis, I proceed by examining the nodes in each region in

more detail. I focus my examination of each region of the graph primarily on trackers associated with the audience marketplace and less on trackers essential to website functioning or to their privacy features.

As I have noted in another article co-authored with several colleagues, by now the iterative nature of such exploratory work should be apparent, as well as the importance of grounding classificatory work into network topology and letting network structure demarcation be informed by classificatory work:

> It is important to notice that the operation of classifying the nodes and of reading the disposition of classes are not separated, but performed at the same time. As it will become clear in the next pages, our technique does not consist simply in the projection of a set of pre-existing categories on a connectivity-based layout, but on recursively using the categories to make sense of the layout and the layout to define the categories. (Venturini et al., 2018, p. 8)

Finally, one classification criterion might not be enough to make sense of the disposition of nodes in a network which depicts a complex empirical phenomenon, and certainly not in the case of bipartite networks where each node type asks to be treated according to its specificity. Hence multiple features of a first-party site or a third-party tracking element might need to be considered to make sense of the topology of the network: from genre and business model in the case of first-party sites, to the more specific role in the advertising industry and the audience marketplace for third-party tracking elements.

5.4.2 Variations in Tracking Styles and Audience Marketplace Configurations

The analysis of the key regions of the graph (see Figure 27) suggests that different forms of digital cultural production have their own infrastructures and practices for measuring, analysing, intensifying and monetising the activities of their users. In doing so it shows that it is not only tracking services

that shape news infrastructures and audience construction as discussed in section 4.2, but that technical artifacts and social arrangements are mutually constitutive, as different types of digital cultural production may develop their own tracking styles and practices.

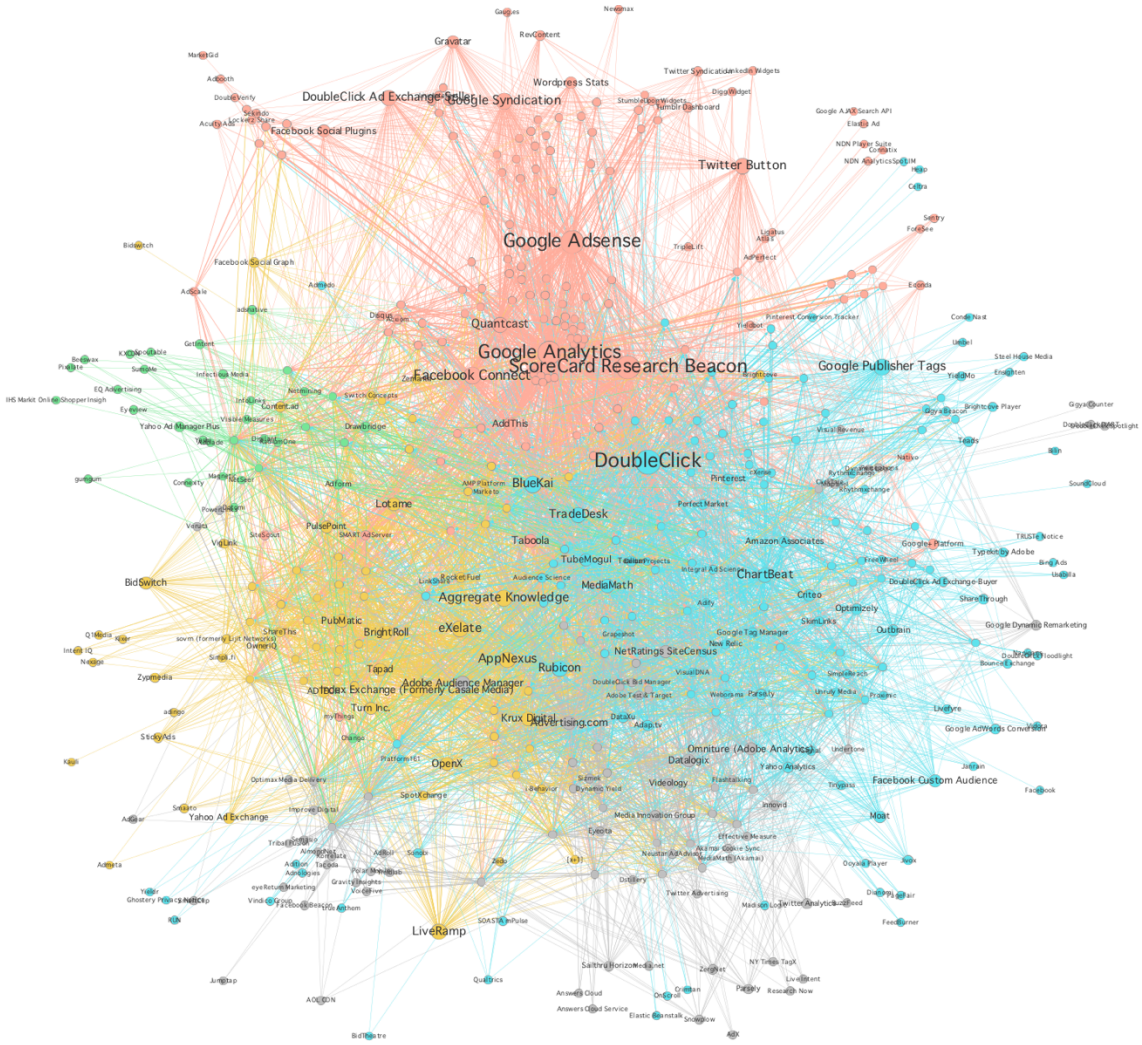


Figure 27: Mainstream and junk news sites and their audience marketplace configurations as seen through third-party tracking networks in which they are embedded. Graph regions or clusters are visually distinguished by colour.

In one of the largest regions of the graph (represented in aqua), we recognise a *professionalised audience marketplace configuration*, associated with high quality publishers (around half of the total number of mainstream news URLs), specialised services and a large number of intermediaries. The audience commodity assembled and exchanged in this space could be imagined as a "premium", highly groomed product, an elaborate construction where users

are known through many data points, collected and combined from multiple sources due to the rather permissive tracking practices of publishers in this space. Indeed, what data points are actually collected by each tracker is certainly an important next step that the type of analysis illustrated in this chapter should explore in future studies.

Audience construction and monetisation in this region is distributed, diversified and intensified through scores of advertising industry intermediaries. These include major ad exchanges such as DoubleClick but also multiple demand-side platforms (e.g. TradeDesk, MediaMath, Madison Logic, TubeMogul and Data Xu). Demand-side platforms are used by advertisers and advertising agencies to automate the buying of targeted inventory from a number of sources. These also include supply-side platforms used by publishers to automate the sale of residual inventory which could not be sold directly to advertisers (such as Rubicon, Teads, Sonobi and Unruly Media). Alongside these we recognise several data brokers or data aggregators specialising in collecting, combining and selling audience data, such as BlueKai, Signal and AudienceScience. Their datasets help to enhance the audience commodity sold to advertisers. In this space audiences are not just measured to sell inventory but also in terms of click-through and conversion rates. This helps to to measure ad effectiveness and is done through services such as Moat, Google AdWords Conversion, and Simple Reach, the latter of which specialises in social media.

In this space we also recognise industry-specific services such as the web analytics startup specialising in news publishers ChartBeat. Audiences are met with content personalisation and recommendation, both of news or marketing, through services such as Optimizely, Parse.ly, Perfect Market and Cxsense. The personalisation of content delivery (both news and ads) is a strategy used to retain visits and increase advertising revenue (Turow, 2011). Paywall system trackers (such as Tinypass) are an indication of the mixed-strategy business models of websites in this cluster. A number of website utilities such as tag or code managers for third-party trackers (Google Tag Manager, Tealium) and website testing and optimisation tools (New Relic), and fonts (Typekit by

Adobe) also indicate a professionalised content production space. Native advertising services (e.g. Outbrain, SkimLinks, ShareThrough) are also present as well as ad retargeting providers (e.g. Criterio), and trackers specialising in the advertising of Amazon products (Amazon Associates).

Compared to the professionalised audience marketplace configuration associated with mainstream news sites, the junk news region of the graph (represented in orange), depicts a *less professionalised audience marketplace configuration*. This configuration is perhaps specific to the long tail of the internet and to small scale, amateur-run publishing operations. The audience construction practices of these operations are also less professionalised and less distributed and seem more rudimentary. In this region we recognise the majority of junk news URLs (over 70% of them). With tracking styles similar to junk news URLs are a couple of mainstream news article URLs belonging to the Guardian and NY Daily News, both reliant on digital advertising for at least part of their funding.

In this space audience monetisation relies on fewer intermediary data and advertising services, and is prominently performed through services that have emerged to monetise user generated content and amateur digital content production or the long tail of the internet. These services include ad networks and ad exchanges, such as Google Adsense, DoubleClick Ad Exchange-Seller, SiteScout and AdScale.

Google AdSense is used by the majority of websites in the network, including by high quality publishers who employ it to monetise their residual inventory or less sell-able audiences, i.e. ad space which has not been directly bought by advertisers. This service uses contextual targeting, in the sense that ads are placed based on a match with a website's content (desilva + phillips, LLC, 2008). In this space we also recognise native advertising networks such as Taboola, RevContent and Nativo. These distribute advertiser sponsored content or "content ads", displayed on websites in the form of recommendations of articles related to those the user is currently viewing. Such services have been criticised for their distribution of false stories and clickbait,

i.e. content with sensational and misleading headlines, on mainstream news sites (Change Advertising, 2016; Griffith, 2014). They have also been criticised for facilitating ad revenue for sites that distribute misleading information (Moses, 2016).

Notable is also the presence of several tracking elements associated with social media, social sharing and social bookmarking platforms. Their presence attests to the platformisation or the expansion of social media audience amplification and measurement mechanisms into junk news spheres, and also to their centrality in attracting traffic to these websites to be monetised through advertising. Such services include Facebook Connect, a service that allows users to connect to third-party websites using their Facebook identity. This service bridges between this junk news region and other regions of the graph. Other social media and social bookmarking services prominent in this region include the Twitter Button, AddThis, Digg, StumbleUpon, LockerzShare, Tumblr, LinkedIn and Google+. The presence of the WordPress blogging platform analytics tracker, Stats, and of the Gravatar service which is integrated with Wordpress also attest to the more amateur character of junk news publishing operations.

Besides Facebook Connect, through their position in the graph we recognise another two audience measurement services in particular as bridging between the different clusters in the network. These are the site-centric traffic analytics service Google Analytics (present in the majority of websites in the network but on more junk news sites than on mainstream news sites) and the cross-site user-centric analytics service ScoreCard Research (also present on the majority of sites), which records audience navigation patterns across websites. The latter is used to generate website rankings which inform decisions about advertising rates and the two are often used in combination (Cardon, 2016). We also recognise the measurement company Quantcast which offers its services to less popular sites which do not perform well in more established website rankings such as Nielsen or comScore (Turow, 2011).

A third large region of the graph (represented in yellow) assembles mainstream

news URLs whose tracking styles appear to resemble at least in part those of junk news URLs. On the junk news side we recognise a mix of satirical and entertainment websites. On the mainstream news side we recognise a public radio network funded through a mix programming fees, grants, sponsorship and advertising, and a political news website funded through subscriptions, advertising and events.

What seems to be specific to audience marketplace configuration in this region is *the prominence of data brokers*, also known as data management platforms or data aggregators. Such services include Aggregate Knowledge, eXelate (owned by Nielsen), LiveRamp and Krux Digital. The role of these data intermediaries is to store, combine, analyse and segment audience data from multiple sources. A particular type stands out, data onboarding services such as LiveRamp, which bring offline data (from customer databases, loyalty programmes or subscriptions), often including personally identifiable information, into online campaigns (Scudder & Wiener, 2012; Joe, 2015). For example, the data collected through Politico's subscription system, one of the mainstream news sites in this cluster, may constitute such a source of offline data. Such data aggregation services inform ad targeting on the advertiser side, and on the publisher side they contribute to the construction of the audience profiles which publishers sell to advertisers (Marvin, 2016).

The audience construction practices of the junk news URLs associated with this audience marketplace configuration may be seen as more professionalised than those of the URLs in the second discussed region of the graph (represented in orange). This is because the use of data aggregators enables publishers to produce more detailed knowledge about audiences to support the sale of inventory. We also recognise video advertising as prominent in this audience marketplace configuration through services such as SpotXchange, StickyAds and Zypmedia. Finally, VigLinks, a content monetisation service that enables the monetisation of referral links, seems to be specific to junk news sites (Khan, 2012).

A final small but notable region (represented in green) consists of a satirical

website (satiratribune.com) with a more *unique tracking profile*. The profile is rather international give its small size, featuring the Russian demand-side platform GetIntent, but also Swiss, Danish and Canadian alongside US services.

Given the fluctuations in tracking practices of junk news sites noted in section 4.3, the network topology discussed above may also be interpreted temporally and indicate distinctions between sites that have been suspended from the services major ad networks and have been pressured to diversify their monetisation strategies.

5.4.3 Asymmetries of Participation in the Audience Marketplace

Another issue that this analysis points towards is that of asymmetries of participation in the audience marketplace. One way in which this is manifested is through concentration or monopolistic tendencies resulting from the platformisation or extension of big platforms in the space of audience measurement and monetisation. This is a tendency which the audience measurement industry has always presented (Napoli, 2011). This phenomenon sees a small number of companies providing services to a large number of websites in the network (on this point see also Libert & Nielsen, 2018). As this analysis has shown, big online platforms such as Google, Facebook and Twitter have come to reshape not just news production and circulation but also audience commodification, and to dominate the online audience marketplace. Three Google services, DoubleClick, Google Analytics and Google Adsense are each present on over 80% of the URLs in the network. Facebook Connect is present on over half of the URLs in the network and the Twitter button on over 40% of them. Another service present on the majority of URLs is ScoreCard Research (over 80%). Although focusing primarily on audience analytics to support editorial decision-making, the presence of the startup ChartBeat alongside the big platforms on over 50% of URLs, is worth noting. Media concentration is also geographically circumscribed around the US.

While the notion of media concentration points towards asymmetries between service providers in the audience marketplace, another asymmetry pertains to the relationship of news publishers and the digital advertising industry. It is striking that advertising revenue has been in decline for news publishers, given the scale at which news sites have become vehicles for intrusive digital advertising and data industry tracking and data collection practices (over two thirds of all tracking elements in this study are advertising-related). For this reason, the tracking practices of mainstream news sites should be addressed not only in relation to user privacy and data protection concerns as initiatives such as the EU General Data Protection Regulation aimed to do, but as Englehardt & Narayanan (2016) also gesture towards, they should also be seen as an indication of the precarity of business models and revenue streams of news organisations which place increasing pressure on monetising content with advertising. While regulations such as the GDPR are aimed at increasing transparency and accountability on the audience - data collector and processor side of the advertising market, the relation between publishers and advertisers and the pressures that the business practices of the online advertising and digital marketing industries place on publishers would also need to be attended to if aggressive online data collection practices and privacy invasion are to be tempered. Indeed, absence of transparency and accountability on this side of the market is of concern as well, and publishers have repeatedly expressed their concern about the low returns they receive from selling online advertising inventory and the lack of transparency and accountability that a highly complex intermediation structure generates (Davies, 2016; Pidgeon, 2016).

Finally, the last aspect of this asymmetry pertains to the publisher-audience relation. While audiences are typically seen to pay for news in two ways, with the time spent reading the news and sometimes by buying access to news content with money (Nielsen, 2016), this analysis reminds us of a third less acknowledged form of payment that audiences make in news, namely their data. While the relationship between news and its publics is typically discussed as citizen participation in democratic societies through news or as audience participation in journalism through interactivity features of digital media (for a

discussion of this see, e.g., Peters & Witschge, 2015), this analysis reminds us that news and its institutions have always been co-produced with publics. Audiences have always actively participated in news through the provision of their data which informs multiple aspects of news, from the financial stability of institutions, to editorial decisions and the authority of a news publication, even if professional claims to journalistic authority do not allow these contributions to be recognised as such.

5.5 Conclusion

This chapter tested the news device approach in the context of another area of news, the marketplaces through which audience products are made and exchanged. In doing so it can be seen as a contribution to socio-material approaches to news work and their calls for bringing such approaches to bear on the business side of news (see, e.g., Lewis & Westlund, 2015). By studying the audience marketplace configurations of news websites, and more specifically of news pages that attract high social media engagement scores, in relation to those of junk viral news sites, it also aims to contribute to debates about fake news and the economics of advertising-supported viral content production.

The approach developed in this chapter explores the capacities of the website as a site for studying the business side of news. More specifically it focuses on the third-party user tracking mechanisms embedded in websites and their participation in the production and exchange of audience products. By combining audience economics with an infrastructural or material approach to new media studies, I illustrated an approach by means of which such digital objects can be used to study the composition and relations that make up the post-exposure audience marketplace configurations in which advertising-supported mainstream news and viral junk news sites operate.

In relation to the central question of this thesis, pertaining to how digital objects participate in and format news work and ways of studying it, this

chapter invited an understanding of how digital devices and their associated industries make a difference to news not only through the shaping of practices but also through the shaping of the infrastructures of news, in this case the website, through the integration of third-party elements.

Another contribution is made by extending the study of interactions between news and digital devices to also include interactions between these devices and knowledge creation. I describe how a tracking detection device, Ghostery, can be configured into a research device with the help of the Tracker Tracker research tool and analytical techniques developed to address the research question of this chapter. Such techniques enrich tracker detection at the level of an individual website with the analysis of associations between tracking elements based on their shared use by websites. I suggest that attention needs to be paid to how the research method shapes the research object and describe how the composition of audience marketplaces is configured by the dynamic character of tracking. In doing so I suggest that the study of audience marketplace configurations cannot be separated from the study of tracking practices, which shape the representations of audience marketplaces we arrive at.

A contribution is made to the study of audience marketplaces by providing an account from the viewpoint of tracking infrastructures of websites. While this may be seen as a partial view, and indeed it is, as any other representation, it may also be seen as enriching ways of knowing audience marketplace practices with another perspective, the view put forward by the tracking infrastructures of websites. From the point of view of tracking infrastructures of websites, junk news, and to a much greater degree mainstream news, appear to be deeply entangled with the complex structures of the online advertising and marketing industries. The economic dependence of news on the online advertising industry is manifested at the level of the material infrastructure of news, the news site, through an invisible tracking infrastructure, through which relations, exchanges and data flows are established between different participants in the online audience marketplace. This dense and invisible tracking infrastructure is problematic not only from the point of view of audiences in relation to issues

of user privacy, security and labour but, as far as news publishers are concerned, it is indicative of the precarity of business models and revenue streams of news organisations which place increasing pressure to sustain resource intensive news production through complex and invasive advertising structures, that rely on aggressive data collection practices, with little transparency and accountability towards other participants in the audience marketplace.

I illustrated a research technique by means of which specificities of tracking and audience marketplace practices can be surfaced and compared across digital cultural production. The analysis suggested that different forms of digital cultural production have their own practices and infrastructures for intensifying, measuring, analysing and monetising the activities of their audiences. This suggests that while the online advertising industry and its tracking infrastructures is shaping not just audience economics but also news infrastructures, tracking infrastructures are also shaped by cultural production dynamics. This study illustrated a few such audience marketplace arrangements, from amateur configurations associated with the internet long tail, to more complex configurations associated with professionalised forms of information production, such as the news media.

Asymmetries between participants in the audience marketplace cut across all these configurations, from monopolistic tendencies of big online platforms in the online advertising industry, to economic pressures on publishers which increasingly become vehicles for the aggressive data collection practices of advertising industry actors, and finally audience members who increasingly bear responsibility for the implications of these market configurations for their privacy and security, aided by data protection software and regulations (such as the recently enforced EU General Data Protection Regulation).

This analysis provided a picture of audience marketplace practices in which legal initiatives such as the recently enforced EU General Data Protection Regulation are aiming to intervene. Future research may use this analysis as a baseline to examine the changes that this regulation has brought to audience

marketplace configurations for users accessing publisher websites from Europe. Finally, while this analysis focused on display advertising, future research should also examine the audience marketplaces associated with mobile news sites as well as apps.

6. Conclusion

In this thesis I set out to empirically explore the interactions between digital devices, news work and journalism research. I started this project imagining that promising new digital methods and device approaches could help make a contribution towards understanding digital transformations of news and associated forms of work.

For this reason, I set out to test and further elaborate what might be entailed in bringing such approaches to bear on the study of news through a series of empirical cases. Such techniques come with challenges and are sometimes met with ambivalence. They can either be suspected to be prone to misrepresenting the studied phenomena due to various digital biases, or they can be embraced as sources of data about social phenomena, without paying attention to how they interfere with them, and treated alongside big data approaches. While both these positions represent traps that digital research can fall into, the research techniques I explore in this dissertation offer means to avoid these problems. But to do so they call on us to slow down, to stay with the troubles we come across (to use Haraway's (2016) language), and let digital devices interfere with the topics we study, the methods we develop and the research problems we address.

My main argument in this dissertation is that the study of digital news and journalism could benefit from approaches that treat news work as co-produced through interactions with digital devices, and that to account for these interactions we can leverage the affordances of digital devices for research. I proposed the notion of news device to capture this twin proposal to attend to the role of digital devices in news work and in digital journalism research.

I argued that how digital objects participate in news work and in journalism research is an important question. Digital devices are habitually used to make and consume news and they are becoming an increasingly salient part of infrastructures that sustain our information spheres. While these devices are increasingly being recognised as a significant participant in news work, news is not necessarily widely recognised yet as a hybrid, socio-material practice, but

rather still imagined to be organised primarily through what Deuze and Witschge (2018) describe as stable and universal professional cultures, occupational norms and routines. As far as digital journalism research is concerned, it is an important question because the possibilities that the digital affords for journalism research are not exhausted by current approaches that treat the digital either as a new object of study for established research methods, as an unmediated source of data, or as computational methods for the study of established practices.

In this final chapter I draw together the key contributions of this dissertation and provide some final reflections on their potential, limitations and implications for the study of digital news and journalism, as well as some research directions that deserve to be explored further.

6.1 Key Contributions

As it should be clear by now, the aim of this thesis is not to provide a general theory of the transformations of journalism in the digital age. Instead, more modestly, this thesis aims to contribute towards the collective endeavour of understanding these transformations in two ways. Firstly, it proposes a research approach to address the interactions between digital objects, news work and journalism research. Secondly, it makes an empirical contribution by testing this approach in the context of three case studies and proposing concepts that help to make sense of some of the aspects of news in the digital age. My research approach and empirical investigations are "theoretically informed" (Becker, 1998), in the sense that, as detailed in Chapter 2, the conceptual underpinnings of socio-technical approaches to the study of journalism, and of device perspectives from digital social and media research, have shaped decisions about what to study and how to conduct research. In what follows, I will summarise each of the contributions of this dissertation.

6.1.1 News Devices: A Device Approach to Digital Journalism Research

This dissertation contributes to news and journalism research by developing a device-sensitive perspective to approaching the digital transformations of news and their implications for journalism research. The news device approach is a term I use to describe the combination of socio-material approaches to journalism with device-centred perspectives from digital social and media research. Inspired by Muniesa et al.'s (2007) notion of market devices, this concept draws attention both to how the digital offers sites, techniques and practices through which news work can play out, but also to how it can offer means to study both news and digital devices.

News device approaches carefully attend to how relations and practices are inscribed, supported and enacted by digital objects. This involves asking not just to how digital devices are used for journalism but also how digital devices treat, process or enact various aspects of news work, how they configure the relations between news and other domains, and, perhaps, what news becomes in the context of digital devices.

Two contributions are made through this approach. First, such a perspective invites an understanding of digital news and journalism as varied socio-material practices situated in and materialised through "fields of devices" (Ruppert et al., 2013), even if this account might not always correspond to news professionals' understanding of their profession. Understanding the "digital" in digital journalism as the proliferation of diverse digital objects in mundane news settings invites us to attend to the specificities of their interactions with news and journalism.

In its empirical orientation towards the questions of when and how devices come to matter in situated practices, this approach can be distinguished from approaches that treat the question of impact of digital technologies on journalism in a monolithic and undifferentiated way, or that seek to address it exclusively on theoretical grounds.

The news device approach guides attention both to how digital devices shape *news* in particular situations, and towards examining how digital devices assemble and materialise *relations* between news and other areas and domains, from other digital content producers, to digital visual culture, commercial online platforms, and the online advertising and marketing industries. From this point of view, this whole project can be read as a contribution to understanding the ongoing practices of "hybridisation" or blending of news with other domains (Chadwick, 2013), even while journalistic discourses may focus on "boundary work" (see, e.g., Carlson & Lewis, 2015).

A second contribution pertains to the proposition that inquiries into the interactions between news and the digital should be extended to also cover the role of digital devices in journalism *research*. To this end, I discuss a set of methodological tactics by means of which digital devices can be configured to support news and journalism research. Following Marres (2017a), this process of configuration should be understood not as the application of well-established protocols but rather as methodological experiments or tests of the capacities of digital devices to inform the research orientations I was interested in. The contribution here is to describe these tactics in action and document the process of configuration in detail, from a description of the analytical capacities of each device and the "conditions of production" (Moats, 2016) of device data through the interactions between technical infrastructure and cultures of use, to corpus demarcation and analytical operations. The tactics I discuss are ones that I have devised or that I adapted from device-centred digital social and media research and are by no means exhaustive.

This second contribution of the news device approach to news and journalism research may also be understood as a way to extend the shift from "social" to "socio-technical" in accounting for news and journalism to the difference that the digital might make to sites and methods for news *research*, rather than trying to minimise their interference. It can also be seen as a contribution to the methodological questions raised by journalism socio-materiality researchers. While by no means the only or the best way to account for the interactions

between news work and digital devices, I hope these methodological tactics will be seen as a useful complement to existing ethnographic approaches that have been fruitfully used by researchers in this area so far.

6.1.2 The News Device Approach in Action: Three Empirical Applications

The main contribution of this dissertation is empirical. It consists in providing insights into the participation of three digital objects in news work and research by putting the device approach to test in the context of three case studies. The case studies took up the intersections between three aspects of news work and three digital objects (see Table 2).

Chapter	3	4	5
Area of news work	Making narrative and storytelling	Making infrastructure and coding	Making audience
Digital object	The network graph	The coding platform	The web tracker
News device	The network as storytelling device	GitHub as connective coding device	The tracker as audience marketplace device
Research technique	Multimodal analysis and graph semiotics	Platform interface and documentation analysis; Extracting and configuring platform metrics and date stamps to study characteristics of journalism coding on GitHub	Configuring web tracking detection on individual sites to examine tracking networks across websites with the Tracker Tracker and visual network exploration
Findings	5 types of network stories co-produced by network material affordances and journalism genre conventions	News work as connective coding; 4 characteristics of platform-specific news coding practices	Multiple tracking styles; Spectrum between amateur and professional audience marketplace configurations

Table 2: Summary of the three empirical applications of the news device approach.

In what follows I will summarise the configuration and findings of these case studies. These cases do not exhaust the many ways in which digital devices

participate in news and research. Rather, they illustrate three salient ways in which they do so and are of course only a handful of the many configurations that can be described as news devices. Moreover, there are of course many actors that participate in situations where digital devices and news intersect. As suggested in the Introduction, in the empirical world news happens through a multitude of intersecting digital objects and devices, but also professional norms and values, and is experienced alongside other sources and types of information. While the case studies have been constructed around one type of digital device, the chapters discuss the contributions of many intersecting devices.

6.1.2.1 Networks as Storytelling Devices

The news device that the first empirical chapter focused on was the network diagram. I focused on this digital object because network visualisations are a landmark of digital visual culture and an increasingly important means to analyse and represent collective phenomena in a number of domains. In the context of journalistic knowledge making they are a less established but growing mode of visual representation. What also makes it interesting is that, unlike the other two digital objects examined, it is not a born digital or natively digital object, in the sense that it is not specific to digital environments but rather pre-exists them (Rogers, 2013).

In the spirit of Boczkowski and Mitchelstein's (2017) suggestion to build bridges with academic communities outside the news research community, the configuration of network diagrams as an object of study in this chapter was informed not just by concerns pertaining to news research but also by broader interests in the narrative potential of quantification practices, and of network visual exploration more specifically. Hence in this chapter network diagrams were problematised as narrative devices and the focus was on how network diagrams are performative of the way in which aspects of collective life are rendered into journalistic stories and the kinds of stories that are being told.

The capacities of networks to make a difference to narratives are approached

through their affordances, i.e. through the capacities of their visual attributes (e.g. node position, size and hue) and the network properties these attributes materialise (e.g. clustering, ego-networks, weak ties), to elicit narrative readings — in combination with other elements of the journalistic story (headline, lead paragraph, graph caption, etc.). What the news device perspective affords in this case is an analysis of narrative from the point of view of a particular storytelling device, the network diagram, which in this chapter I capture with the notion of network stories. I show how narrative readings are co-produced through the affordances of network graphs, the journalistic genres they are embedded in and the reader's own socio-cultural knowledge.

This perspective results in insights about meaning making around journalistic network stories. I find that there are recurring ways in which meaning is construed out of journalistic network stories which include exploring associations around single actors, detecting key players, mapping alliances and oppositions, exploring the evolution of associations over time, and revealing hidden ties. Multiple of these narrative readings can be encountered in a journalistic piece, particularly when these pieces include interactive network diagrams. While neither representative nor comprehensive, these can be seen as a contribution towards a vocabulary of narrative readings of networks in journalism as well as towards a protocol for the construal of narrative meaning out of networks.

The analytical approach I use in this chapter is different from those developed in the next two chapters, although there are also overlaps. The analysis developed in this chapter draws on well-established social-semiotic approaches to the study of meaning making in multimodal communicative texts. In doing so I aim to build a bridge towards existing analytical approaches and to show that these are well suited for the analysis of news devices. But given that multimodal analysis is a well-tested, well-established and well-documented approach, I draw on it to illustrate the construction of networks as narrative devices but do not make it a central part of the research problem that this thesis addresses. Instead, I focus the problematisation of this dissertation on less established but promising device-centred analytical approaches to the

digital. These bring new challenges that require more testing and elucidation before they can become useful ways of approaching digital news and journalism.

On the other hand, multimodal analysis and research tactics associated with device-centred approaches are not completely distinct. What this chapter shares with Chapter 5 is the presence of the network diagram. While Chapter 3 takes the network diagram as an object of analysis, the network diagram becomes part of the visual network exploration technique I use in Chapter 5. In the first case we are dealing with the use of network diagrams for communication purposes, while in the second one we are dealing with the use of networks for exploratory analysis. The twin problematisation of the network visualisation as object of study and method is thus realised across the two chapters. Moreover, multimodal analytical approaches used in this chapter are also present in the research techniques used in Chapter 5 as visual network exploration incorporates the multimodal construction and reading of networks. The distinctions between natively digital and pre-digital methods thus are blurred as natively digital approaches are co-extensive with and incorporate established methods in their assembly.

6.1.2.2 GitHub as Connective Coding Device

The second news device that I focus on is the online platform, and more specifically a code sharing platform and its code repositories. I focus on an online platform because they have become important actors in today's news media. GitHub is the largest code sharing platform and one of the most used for news work. GitHub's participation in news is particularly important to examine due to the specificity of this platform as a site for digital infrastructure making (Mackenzie, 2018), including that of news. GitHub was also relevant to study in the context of great interest from the news research community in the role of programming and open source software in news work.

The configuration of the device as an object of study in this chapter does not revolve around how the platform is used by journalists for open source coding,

but around the equally important question of how journalism coding is structured through GitHub. The question is informed by platform and software studies, and their sensitivity towards the conditions of possibility that software enables and software's relation with economic imperatives of platforms.

This analytical orientation allows us to address the question of the changing nature of news in the digital era by contributing to the understanding of one particular process, namely that of the platformisation of news. I provide new insights to this area of news transformations by examining the platformisation of journalism coding. Through user interface and platform documentation analysis, I show how GitHub platformises the coding work that makes use of the platform, i.e. how it turns these forms of work into productive parts of the platform ecosystem through processes of tracing, counting, calculating, recommending, intensifying, multiplying, archiving and mining participation. I call these forms of work connective coding to draw attention to the particular way in which journalism coding is co-produced with the platform. Connective coding expands the understanding of social coding beyond the connectedness or networking functionalities that GitHub enables to also capture the conversion of public coding, developer profiles and behaviours into assets that have the potential to be variously capitalised by the platform and its ecosystem.

The understanding of how the platform structures coding is also important for the configuration of the platform as a research device. Indeed, the same methods through which the platform formats, monitors, networks, ranks and metricises code repositories and user accounts, may afford modes of studying how journalism initiatives inhabit the platform.

Hence the second aspect of the news device approach that this chapter probes is how networked code repositories can be configured to enable the analysis of a collection of news code repositories on GitHub, with a particular focus on surfacing characteristics of platform-specific news coding practice. This has not been an easy task. A number of operations are required on the side of the researcher to align the analytical capacities of the platform with the research

question and there is a constant risk of mirroring the dominant analytical modes of the platform (what is popular and trending), which is a less insightful outcome. I illustrated how researchers can take advantage of the platform's research affordances while at the same time push back against them when they do not align with their own research questions and interests.

I experimented with configuring the capacities of platform metrics and date stamps to tell something about types of platform work, temporality of coding work and how journalism code is valued on GitHub. A number of insights emerged from these analyses.

As far as platform specific types of work are concerned, an analysis of the use of the forking function in my collection of repositories showed that journalism initiatives participate in the platform's code ecosystem with original work at a greater rate than numbers reported by other studies for the entire platform, while also engaging in imitative work. As far as temporality is concerned, many journalism repositories are largely ephemeral and do not conform with the platform's update culture. As far as how journalism code is valued on GitHub, I showed that the recursive nature of the platform's dominant programming publics shapes the valuation of domain-specific journalistic production and steers it towards materials that sustain developers and their software development work. Domain specificity in this case is reflected in the high valuation and ranking of non-code repositories addressed to journalists and non-programming publics, such as repositories containing datasets or data editorial and analytical guidelines. This analysis further nuances the characterisation of platform practices as connective coding by showing that, just as platforms are not fixed or stable arrangements, social practices are also not uniform materials to be platformised but rather are fluctuating, diverse and variable. This mix of modes of valuation reflected in the top starred journalism repositories attests to how, by entering the platform ecosystem, journalism objects may open themselves up to other modes of valuation.

6.1.2.3 Trackers as Audience Marketplace Devices

The news device that the final empirical chapter focuses on is web trackers. Web tracking techniques and associated data mining practices are controversial and have been studied in relation to a number of issues, from surveillance to security, web economies and digital labour. In this chapter web trackers are configured as news devices to explore the role that they play in the business side of advertising-supported news, and more specifically in the making of audience products. I focus on web trackers because in the context of the post-exposure audience marketplace, digital objects such as cookies and other web tracking devices play an increasingly prominent role. My interest in these objects was also prompted by the fake news scandal and associated debates about the economics of junk viral content production, which is why in this chapter I trace the tracking practices of a small corpus of mainstream and junk news sites active around the 2016 US presidential elections.

I examine the tracking infrastructures of these websites to understand what they can bring to our knowledge about the audience marketplaces in which various forms of digital cultural production operate. Audience marketplace configurations in this context can be understood as assemblages of actors that variously participate in audience commodification through web tracking.

To show what a material-empirical approach can bring to the study of media audience commodification I use a technique that configures the tracker detection and classification capacities of the popular Ghostery privacy protection browser extension with visual network exploration. I document the difficulties that come with such an analysis due to the instability of the object of study and the dynamic character of tracking, as well as how the technique is shaping the picture of the phenomenon obtained.

I qualitatively explore post-exposure audience marketplace configurations as materialised through the invisible tracking infrastructures of sites associated with two forms of advertising-supported digital cultural production. A number of insights about web tracking practices and the relations through which the

business side of news plays out can be taken from this analysis. From the point of view of the tracking infrastructures of these websites, junk news, and to a much greater degree mainstream news, are deeply entangled with the complex structures of the online advertising and marketing industries. This tracking infrastructure is indicative of the precarity of business models and revenue streams of news organisations, which places increasing pressure to sustain resource intensive news production through complex and invasive advertising structures, that rely on aggressive data collection practices, with little transparency and accountability towards other participants in the audience marketplace.

It is not only the online advertising and marketing industries that shape the digital infrastructures of news, but tracking practices and infrastructures are also shaped by modes of cultural production. Different forms of digital cultural production have their own practices and infrastructures for intensifying, measuring, analysing and monetising the activities of their audiences. This study illustrated several such audience marketplace configurations, placed on a spectrum from amateur audience marketplace configurations specific to the long tail of the internet, more moderate in tracking and heavily reliant on social media to attract traffic and ad networks to monetize it, to professionalized configurations where tracking is intensified and customised through specialised services and large numbers of media buying and selling intermediaries, as well as data brokers.

Finally, asymmetries between participants in the audience marketplace cut across all these configurations, from monopolistic tendencies of big online platforms in the online advertising industry, to economic pressures on publishers which increasingly become vehicles for the aggressive data collection practices of advertising industry actors, and finally audience members who increasingly bear responsibility for the implications of these market configurations for their privacy and security, aided by data protection software and regulations (such as the recently enforced EU General Data Protection Regulation).

6.2 Implications for the Study of News and Journalism

In this section, I offer reflections on what the contributions discussed in the previous section might mean for the study of digital news and journalism more generally, as well as for the areas of journalism studies and digital social and media research. While I do not want to generalise too much from my three case studies, I will add a few considerations on the potential and challenges that come with these approaches based on my empirical research.

Digital media, online platforms and the web, for short, the digital, are present in journalism research mainly as either a research topic or as a source of data about news and news work. The contribution of this dissertation has been to outline a way to approach digital devices at once as research objects and as being able to be configured into device-specific modes of knowing the interactions between news and digital devices.

This is not the attractive promise of big data to deliver unmediated access to large amounts of granular data about entire populations, and computational techniques that would enable large-scale analyses of these datasets. In this dissertation I illustrated that the promise of device-driven research perspectives consists in the pairing of critical research with configuring the analytical modes inscribed in devices to understand their interactions with issues and practices. This is surely a more modest and, to some, less attractive and more difficult promise. But what makes this approach difficult is also where its potential lies.

A news device approach can make available new sites and research techniques to address key questions about news. But the conceptual and methodological outlooks that underpin it, whether that is socio-material approaches, device-driven research, digital methods, platform studies or software studies, will also modify these questions, objects of study and research problems. By treating the digital device not just as a collection of data but also as an object of investigation, the ambiguity about the extent to which you are studying news practices or device effects (present, for example in Chapter 4 concerned with platform specific journalism coding), has implications for how the object of

news and journalism research is conceived. The conceptual and methodological underpinnings of this approach may also introduce different perspectives on what matters in the interactions between digital devices and news, and thus generate possibilities to ask fresh questions and raise new research problems. For example, one common modification when it comes to research questions pertains to the switch from use-centred questions (i.e., how a given device is used in a given practice), to materiality sensitive questions of how a device treats or configures a practice or an issue. This was the case for example in Chapter 4 where I decided not to focus on the question of how journalists use GitHub to develop open source software, which could potentially be seen as better aligned with the commitments of journalism studies, and instead focused on how the platform structures journalism coding. While this modification may seem trivial, this shift is intended to re-focalise the research perspective to account for agency not as solely the domain of the users but as being distributed between users and platform. This in turn has implications for the operationalisation of the question, i.e. for how the different elements of the research apparatus are aligned to address question.

This interference, which I see as positive, of the conceptual and methodological outlooks that underpin the news device approach with objects of study and research problems in news research, may also be seen as a way to set up "two-way streets" between the journalism research community and other research communities, as Boczkowski and Mitchelstein (2017) have suggested. This dissertation illustrated an opening up of news and journalism research towards materiality-sensitive new media studies and digital sociology. I am not arguing that journalism researchers should become new media researchers or digital social researchers, but rather suggesting that one possible direction that the study of cross-media news work might further explore is the potential of device-centred perspectives from digital social and media research, software studies and platform studies, as several journalism researchers have already begun to do.

As mentioned above, news device approaches come with many challenges, and difficult ones to resolve for researchers used to working within more

established methodological frameworks. Embracing these approaches would also involve embracing particular ways of addressing these challenges, which may further unsettle how news and journalism are studied. I will briefly discuss some of these challenges below. I will not aim to reassure researchers by providing principles through which each of these can be safely averted or overcome once and for all. As I suggested above, these research approaches do not necessarily aim to remove the interference of the device with the object of study or with the data. The solution they propose is to modify the research problematisation and make them part of the topic to be investigated. Indeed, what constitutes a problem depends on how the research apparatus is configured (for a good discussion on this point, see Weltevrede, 2016).

One risk that the case of studying journalism coding on GitHub surfaced was that the reliance on "methods of the medium" would pull the research in the direction of reproducing the modes of analysis that the platform offers (e.g. trending and popular content and influential users), without modifying them towards the understanding of aspects of platform-specific coding practice. Whether this pull is a good or a bad thing ultimately depends on the research question and the objective of the research. In any case, a number of research techniques are available to researchers to align the analytical affordances of devices with the questions of social and media research (see, e.g., Marres, 2017a; Rogers, 2013; Weltevrede, 2016). Researchers are also encouraged to develop their own, as every device will require its own research techniques. In my empirical studies, I relied on a handful of approaches developed in the context of previous research, such as the move from frequency to relational analyses which I used in Chapter 5 (for more on this technique, see, e.g., Marres & Gerlitz, 2015). In sum, one key to mitigating this risk, as Weltevrede (2016) suggests, stands in the "quality of configuration" of the research apparatus through the alignment of questions, with data and research techniques.

Another reason for hesitation concerns issues of data access and data collection. The reliance on platform APIs (as has been the case in Chapter 4) opens up questions about what data is made available and how data is

structured via APIs. While what is available through APIs and the terms of use vary from platform to platform, APIs have politics, as Puschmann and Burgess (2013) suggest. More recently, there has been talk of a possible crisis of API research in social media studies, as platforms are restricting access to data and social datasets are increasingly put into question following recent scandals around fake news and misinformation (see, e.g., Rogers, 2018). My approach when it comes to accessing data via the GitHub API in Chapter 4, following Marres and Weltevrede (2013), has been to make the affordances of this API and how it structures knowledge making, a topic of investigation. Investigating and not abandoning APIs is increasingly important as journalists themselves are using APIs in their work more and more. Similarly, the reliance on a commercial tool such as Ghostery in Chapter 5 can cause hesitation. The researcher is dependent on the tracker detection techniques of the tool which are tied to the service's own goals. This is not an easy problem to solve but, as I suggest in my chapter, it can be mitigated by making the construction of tracking detection through this tool part of the research object and the findings.

Linked to these are also concerns about the opacity and the constantly changing nature of the often algorithmic processes that underpin digital devices. In the case of my research the issue of opacity came up in the case of the GitHub trends algorithm which plays a role in shaping engagement with code repositories. While we know what platform activities the algorithm takes into account, we do not know exactly how they are configured and what weight they are given. The instability of digital devices surfaced in relation to the tracker detection service used in Chapter 5. For example, Ghostery has changed the classification of trackers it detects at various points during this research. But device-centred approaches do not abandon digital devices because of these challenges but instead seek to make the configuration of the research apparatus sensitive to these characteristics so as to be able to study them. Dealing with these requires these approaches to be flexible and adaptable both to the shifting nature of the device and in terms of their object of inquiry in order to allow these "epistemic trouble[s]" to become part of the investigation (Marres and Weltevrede, 2013; on device-driven research

approaches as flexible and adaptable see also Weltevrede, 2016 and Niederer, 2016). As far as opacity is concerned, even if, as proprietary "coded object[s]" algorithms might not be available to us for inspection, Bucher (2018) argues that there are still many ways in which we can make sense of them. This is because algorithms are not just proprietary code but many different things depending on the configurations that they enter, and hence many different methodological tactics can be devised to capture multiple aspects of them (Bucher, 2018, p.150), even if we may never know them exhaustively.

Bucher's argument brings to what may be seen as another difficulty, namely the always partial nature of the accounts we produce. This dissertation provides multiple partial representations of the interactions between digital devices and news work: the role of digital devices in making narrative, the role of digital devices in making news infrastructure, and their role in making audience products. The particularities of these partial accounts are shaped by the device perspective and the various approaches used to treat these interactions: reading the role of the device from the content of news texts, reading the role of the device from how it organises code repositories, reading journalism coding from its platform-specific networked character, and reading audience making practices from the point of view of tracking devices embedded in websites.

In this dissertation I did not strive for one response or one approach that would enable me to tame the question of the impact of the digital on news and journalism, and the direction that transformations are taking, in all its complexity, once and for all. I preferred the more modest approach of multiplying partial accounts of particular interactions and describing the different operations through which digital objects come to matter in relation to various aspects of news work. This is akin to what in the context of controversy mapping has been called "second-degree objectivity", which is obtained through the "multiplication of different viewpoints", "from diversity rather than from uniformity" (Venturini & Munk, forthcoming, p. 177; see also Venturini, 2012).

These evaluations are not only partial and particular, but also not definitive given the ever-shifting nature of devices, which is why the emphasis in this dissertation has not been only on the evaluations of their capacities to shape practices and relations but also on describing the means by which others can configure their own assessments.[70]

6.3 Thinking Ahead: Participatory Approaches in News Device Research

Thinking ahead, there are of course many more devices and methodological tactics that could be discussed under the rubric of news devices and associated research approaches. As digital devices evolve, so will device-centred modes of studying them continue to develop. There are also many ways in which the explorations I developed in my empirical chapters could be extended and improved. I offer a few suggestions for future research in the chapters themselves. I will not revisit these in this section and will instead take this final section as an opportunity to more explicitly articulate some of the recurring but hitherto underdeveloped threads that ran through this dissertation into possible future research directions.

As I conclude this dissertation, we find ourselves in a particularly crucial moment for inquiries into the interactions between news and digital devices, as the debates of recent years about the role of devices such as platforms in opinion manipulation, misinformation and the weakening of news institutions, demand increased scrutiny and evaluation.

The approach suggested in this dissertation was that the critical interrogation of digital devices and their implications for news practices and relations does not need to be separated from the use of analytical affordances of digital devices and computation, but that, instead, they can be productively combined.

[70] To this end, some of the research in this dissertation (Chapter 5) has in earlier versions been published as "methodological recipes" that would enable others to investigate the phenomenon of "fake news", as part of *A Field Guide to "Fake News" and Other Information Disorders* (Bounegru et al., 2018).

From this point of view, this is a particularly fruitful moment. This is because critical interventions around the digital that combine interrogation of digital devices with configuring their analytical capacities to serve these inquiries, are increasingly occurring not only in academic settings but in many other areas of society as well, from art to journalism to everyday life. Journalism is one particularly important area where these inquiries have been thriving over the past years, in practices such as algorithmic accountability reporting (Diakopoulos, 2015) and data journalism (Gray & Bounegru, forthcoming).

Elsewhere colleagues and I proposed the notion of "data infrastructure literacy" to draw attention to such practices that focus not just on the ability to work with datasets, but also on the ability to use datasets to critically interrogate and intervene around the socio-technical arrangements through which data is produced and manipulated (Gray et al., 2018). We discussed examples from digital social and media research similar to the kinds of interventions discussed in this dissertation, but also from data journalism and data activism. In data journalism and algorithmic accountability reporting, there are numerous projects that critically and tactically investigate, challenge and mobilise datasets, data infrastructures and algorithmic processes to intervene in defining the fields of action and possibility enabled by digital infrastructures and devices, from double-voter detection systems, to migrant deaths data collection systems, and criminal re-offence risk assessment tools. A look at ProPublica's Machine Bias series will provide many other examples.[71] What all the examples discussed in our research shared, was working to inventively align the analytical affordances of data infrastructures with the aims of critical interrogation, challenging and intervening in the composition of the digital infrastructures that permeate our lives and professional practices.

Thinking ahead, given the importance of the task at hand and the commitment of many stakeholders, including from journalism, in the implications of digital devices for society, participatory approaches to inquiries into and with digital devices might be a direction worth exploring further. In this respect, participatory, engaged, experimental and creative research approaches from

71 https://www.propublica.org/series/machine-bias/p3

STS could serve as inspiration (see, e.g., Lezaun, Marres, & Tironi, 2016; Marres, Guggenheim, & Wilkie, 2018; Chilvers & Kearnes, 2015; Sismondo, 2008). Unlike approaches that demand such research to take "critical distance" from the studied phenomenon to truly understand its implications, participatory approaches advocate for "critical proximity" (Birkbak, Petersen, Elgaard Jensen, 2015; Latour, 2005a). Such critical proximity would see collaborations set up with various stakeholders, such as journalists, to configure inquiries into and with digital devices. These inquiries would take advantage of problematisations already developed in everyday life and in their own professional practices (on this point see also Marres, 2017a; Marres et al., 2018).

Elsewhere, I experimented with this approach by bringing together data journalists and journalism researchers to develop accounts of how they integrate critical interrogation of datafication with data work in their day to day practices, and how such practices may be modified towards what, following Agre, we "critical data practice" (Gray & Bounegru, forthcoming). The news device approach as developed in some parts of this dissertation may be understood as participatory in two ways. First, in a sense, as co-developed with the participation of digital devices, as research questions and their operationalisation are partly informed by the operations of digital devices. Secondly, efforts towards configuring inquiries with actors other than digital devices have been present in Chapter 5 where the problematisation of fake news from an economic perspective has been informed by journalistic investigations into the topic, such as the work of BuzzFeed News on this issue, and versions of this work have been published as journalistic investigations.[72]

The device perspective offers ways of examining, exploring and experimenting with the role of the digital in news and journalism work and research. Rather than treating the digital as a monolithic development with unified effects, it offers a way to look at the mutual articulation between devices and practices in particular settings. Rather than looking at the analytical capacities of digital data from these devices either as biased or as giving new unmediated access to

72 https://www.buzzfeednews.com/article/craigsilverman/fake-news-real-ads

social interactions through big data, the device perspective opens up space for reflecting on what it means for digital devices and data to articulate concerns, issues and practices related to news and journalism in particular ways. I hope the approaches discussed in this dissertation may be taken as an invitation to explore how digital technologies are involved in modifying relations and practices in news and news research in a participatory manner, including in collaboration with various publics who are affected by such changes.

References

Adamic, L. A., & Adar, E. (2003). Friends and Neighbors on the Web. *Social Networks*, *25*(3), 211–230. https://doi.org/10.1016/S0378-8733(03)00009-1

Ahva, L., & Steensen, S. (2017). Deconstructing Digital Journalism Studies. In B. Franklin & S. I. Eldridge (Eds.), *The Routledge Companion to Digital Journalism Studies*. London: Routledge.

Alshwaikh, J. (2009). Diagrams as Interaction: The Interpersonal (meta)function of Geometrical Diagrams. In M. Joubert (Ed.), *Proceedings of the British Society for Research into Learning Mathematics* (Vol. 29(1)).

Amoore, L., & Piotukh, V. (2016). Introduction. In L. Amoore & V. Piotukh (Eds.), *Algorithmic Life: Calculative Devices in the Age of Big Data*. London: Routledge.

Amsterdamska, O. (1990). Book Review: Surely You Are Joking, Monsieur Latour! Science in Action, by Bruno Latour. Milton Keynes: Open University Press: 1987, 274 pp. $25.00. Also available in paper from Harvard University Press, $12.95. *Science, Technology, & Human Values*, *15*(4), 495–504. https://doi.org/10.1177/016224399001500407

Ananny, M. (2013). Press-Public Collaboration as Infrastructure: Tracing News Organizations and Programming Publics in Application Programming Interfaces. *American Behavioral Scientist*, *57*(5), 623–642. https://doi.org/10.1177/0002764212469363

Anderson, C. W. (2011a). Between Creative and Quantified Audiences: Web Metrics and Changing Patterns of Newswork in Local Us Newsrooms. *Journalism: Theory, Practice & Criticism*, *12*(5), 550–566. https://doi.org/10.1177/1464884911402451

Anderson, C. W. (2011b). Blowing up the Newsroom Ethnography in an Age of Distributed Journalism. In D. Domingo & C. Paterson (Eds.), *Making Online News: Newsroom Ethnographies in the Second Decade of Internet Journalism* (Vol. 2, pp. 151–160). New York: Peter Lang Publishing. Retrieved from https://www.peterlang.com/view/product/28752

Anderson, C. W. (2015). Between the Unique and the Pattern. *Digital Journalism*, *3*(3), 349–363.

https://doi.org/10.1080/21670811.2014.976407

Anderson, C. W. (2016). News Ecosystems. In T. Witschge, C. W. Anderson, D. Domingo, & A. Hermida (Eds.), *The SAGE Handbook of Digital Journalism* (pp. 410–423). London: SAGE Publications. https://doi.org/10.4135/9781473957909

Anderson, C. W. (2018). *Apostles of Certainty: Data Journalism and the Politics of Doubt*. Oxford, New York: Oxford University Press.

Anderson, C. W., & de Maeyer, J. (2015). Objects of journalism and the news. *Journalism*, *16*(1), 3–9. https://doi.org/10.1177/1464884914545728

Anderson, C. W., & Kreiss, D. (2013). Black Boxes as Capacities for and Constraints on Action: Electoral Politics, Journalism, and Devices of Representation. *Qualitative Sociology*, *36*(4), 365–382.

Andrejevic, M. (2014). Big Data, Big Questions | The Big Data Divide. *International Journal of Communication*, *8*, 1673–1689. Retrieved from https://ijoc.org/index.php/ijoc/article/view/2161

Andris, C., Lee, D., Hamilton, M. J., Martino, M., Gunning, C. E., & Selden, J. A. (2015). The Rise of Partisanship and Super-Cooperators in the U.S. House of Representatives. *PLOS ONE*, *10*(4), e0123507. https://doi.org/10.1371/journal.pone.0123507

Asay, M. (2018, July 13). Pretty Much No One Quit GitHub Over the Microsoft Acquisition: Here's Why. *TechRepublic*. Retrieved from https://www.techrepublic.com/article/pretty-much-no-one-quit-github-over-the-microsoft-acquisition-heres-why/

Atton, C. (Ed.). (2015). *The Routledge Companion to Alternative and Community Media*. Routledge.

Bach, B., Kerracher, N., Hall, K. W., Carpendale, S., Kennedy, J., & Riche, N. H. (2016). Telling Stories about Dynamic Networks with Graph Comics. Presented at the Proceedings of the Conference on Human Factors in Information Systems (CHI), ACM, New York, United States. https://doi.org/10.1145/2858036.2858387

Badouard, R., Mabi, C., Mattozzi, A., & Schubert, C. (2016). STS and Media Studies: Alternative Paths in Different Countries. *TECNOSCIENZA Italian Journal of Science & Technology Studies*, 22.

Bakir, V., & McStay, A. (2018). Fake News and The Economy of Emotions.

Digital Journalism, *6*(2), 154–175. https://doi.org/10.1080/21670811.2017.1345645

Bal, M., & Lewin, J. E. (1983). The Narrating and the Focalizing: A Theory of the Agents in Narrative. *Style*, *17*(2), 234–269. Retrieved from https://www.jstor.org/stable/42945469

Barabási, A.-L. (2002). *Linked: How Everything Is Connected to Everything Else and What It Means for Business, Science, and Everyday Life*. New York: Plume.

Barocas, S., & Selbst, A. (2014). Big Data's Disparate Impact. *California Law Review*, *104*(3), 671–732.

Bastide, F. (1990). The Iconography of Scientific Texts: Principles of Analysis. In M. Lynch & S. Woolgar (Eds.), *Representation in Scientific Practice*. Cambridge, MA: MIT Press.

Bateman, J., Wildfeuer, J., & Hiippala, T. (2017). *Multimodality: Foundations, Research and Analysis – A Problem-Oriented Introduction*. Berlin: De Gruyter Mouton.

Bateman, John A. (2011). The Decomposability of Semiotic Modes. In K. O'Halloran & B. A. Smith (Eds.), *Multimodal Studies: Exploring Issues and Domains* (pp. 17–38). London: Routledge.

BBC. (n.d.). How Does the BBC Use Cookies for Advertising? - Using the BBC. Retrieved 24 January 2019, from http://www.bbc.co.uk/usingthebbc/cookies/how-does-the-bbc-use-cookies-for-advertising/

Bearman, P. S., & Stovel, K. (2000). Becoming a Nazi: A Model for Narrative Networks. *Poetics*, *27*(2), 69–90. https://doi.org/10.1016/S0304-422X(99)00022-4

Beaulieu, A., Scharnhorst, A., & Wouters, P. (2007). Not Another Case Study: A Middle-Range Interrogation of Ethnographic Case Studies in the Exploration of E-Science. *Science, Technology, & Human Values*, *32*(6), 672–692. https://doi.org/10.1177/0162243907306188

Becker, H. S. (1982). *Art Worlds*. Berkeley: University of California Press.

Bell, A. (2005). News Stories as Narratives. In I. Mani, J. Pustejovsky, & R. Gaizauskas (Eds.), *The Language of Time: A Reader* (pp. 397–409). Oxford: Oxford University Press.

Bell, A. (2007). The Discourse Structure of News. In T. A. van Dijk (Ed.),

Discourse Studies: A Multidisciplinary Introduction: Discourse as Structure and Process (pp. 58–89). London: SAGE Publications.

Berry, D. (2011, September 12). Messianic Media: Notes on the Real-Time Stream. Retrieved 17 April 2018, from http://stunlaw.blogspot.com/2011/09/messianic-media-notes-on-real-time.html

Bertin, J. (1983). *Semiology of graphics: diagrams, networks, maps*. Madison, WI: University of Wisconsin Press.

Biazzini, M., & Baudry, B. (2014). 'May the Fork Be with You': Novel Metrics to Analyze Collaboration on Github. In *Proceedings of the 5th International Workshop on Emerging Trends in Software Metrics (WETSoM 2014)* (pp. 37–43). Hyderabad, India. Retrieved from https://dl.acm.org/citation.cfm?id=2593875

Binns, R., Lyngs, U., Van Kleek, M., Zhao, J., Libert, T., & Shadbolt, N. (2018). Third Party Tracking in the Mobile Ecosystem. In *Proceedings of the 10th ACM Conference on Web Science* (pp. 23–31). New York, NY, USA: ACM. https://doi.org/10.1145/3201064.3201089

Birkbak, A. (2016). *Caring for Publics: How Media Contribute to Issue Politics*. Aalborg University, Denmark. Retrieved from http://vbn.aau.dk/da/publications/caring-for-publics(f05b27d9-f2c5-45f2-a962-80bf059fd481).html

Birkbak, A., Petersen, M. K., Elgaard Jensen, T., & Philosophy Documentation Center. (2015). Critical Proximity as a Methodological Move in Techno-Anthropology: *Techné: Research in Philosophy and Technology*, *19*(2), 266–290. https://doi.org/10.5840/techne201591138

Blumler, J. G. (2010). Foreword: The Two-Legged Crisis of Journalism. *Journalism Practice*, *4*(3), 243–245. https://doi.org/10.1080/17512781003802576

Boczkowski, P. J. (2004a). *Digitizing the News: Innovation in Online Newspapers*. Cambridge, MA: MIT Press.

Boczkowski, P. J. (2004b). The Processes of Adopting Multimedia and Interactivity in Three Online Newsrooms. *Journal of Communication*, *54*(2), 197–213. https://doi.org/10.1111/j.1460-2466.2004.tb02624.x

Boczkowski, P. J. (2015). The Material Turn in the Study of Journalism: Some

Hopeful and Cautionary Remarks from an Early Explorer. *Journalism*, *16*(1), 65–68. https://doi.org/10.1177/1464884914545734

Boczkowski, P. J., & Lievrouw, L. A. (2008). Bridging STS and Communication Studies. Scholarship on Media and Information Technologies. In Edward J. Hackett, Olga Amsterdamska, Michael Lynch, & Judy Wajcman (Eds.), *The Handbook of Science and Technology Studies (*Third Edition, pp. 949–977). Cambridge, MA: MIT Press.

Boczkowski, P. J., & Mitchelstein, E. (2017). Scholarship on Online Journalism: Roads Traveled and Pathways Ahead. In C. W. Anderson & P. J. Boczkowski (Eds.), *Remaking the News: Essays on the Future of Journalism Scholarship in the Digital Age*. Cambridge, MA: MIT Press.

Boczkowski, P. J., & Siles, I. (2014). Steps Toward Cosmopolitanism in the Study of Media Technologies: Integrating Scholarship on Production, Consumption, Materiality, and Content. In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds.), *Media Technologies: Essays on Communication, Materiality, and Society* (pp. 53–75). Cambridge, MA: MIT Press. Retrieved from http://mitpress.universitypressscholarship.com/view/10.7551/mitpress/9780262525374.001.0001/upso-9780262525374-chapter-4

Bødker, H., & Brügger, N. (2018). The Shifting Temporalities of Online News: The Guardian's Website from 1996 to 2015 , the Shifting Temporalities of Online News: The Guardian's Website from 1996 to 2015. *Journalism*, *19*(1), 56–74. https://doi.org/10.1177/1464884916689153

Bogost, I., & Montfort, N. (2007). New Media as Material Constraint an Introduction to Platform Studies. Presented at the 1st international HASTAC conference, Duke University, Durham NC. Retrieved from http://bogost.com/downloads/Bogost%20Montfort%20HASTAC.pdf

Bogost, I., & Montfort, N. (2009). Platform Studies: Frequently Questioned Answers. Presented at the Digital Arts and Cultures 2009, Irvine, CA. Retrieved from http://nickm.com/if/bogost_montfort_dac_2009.pdf

Borger, M., Hoof, A. van, Meijer, I. C., & Sanders, J. (2013). Constructing Participatory Journalism as a Scholarly Object. *Digital Journalism*, *1*(1), 117–134. https://doi.org/10.1080/21670811.2012.740267

Boumans, J. W., & Trilling, D. (2016). Taking Stock of the Toolkit: An Overview of Relevant Automated Content Analysis Approaches and Techniques for Digital Journalism Scholars. *Digital Journalism*, *4*(1), 8–23. https://doi.org/10.1080/21670811.2015.1096598

Bounegru, L., Gray, J., Venturini, T., & Mauri, M. (2018). *A Field Guide to 'Fake News' and Other Information Disorders*. Amsterdam: Public Data Lab. Retrieved from https://papers.ssrn.com/abstract=3097666

Bounegru, L., Venturini, T., Gray, J., & Jacomy, M. (2017). Narrating Networks: Exploring the Affordances of Networks as Storytelling Devices in Journalism. *Digital Journalism*, *5*(6), 699–730. https://doi.org/10.1080/21670811.2016.1186497

Braun, J. A. (2015). News Programs: Designing Msnbc.com's Online Interfaces. *Journalism*, *16*(1), 27–43. https://doi.org/10.1177/1464884914545730

Broersma, M., & Graham, T. (2013). Twitter as a News Source. *Journalism Practice*, *7*(4), 446–464. https://doi.org/10.1080/17512786.2013.802481

Brotherton, E. A. (2012). Big Brother Gets A Makeover: Behavioral Targeting and the Third-Party Doctrine | Emory University School of Law | Atlanta, GA. Retrieved 2 July 2018, from http://law.emory.edu/elj/content/volume-61/issue-3/comments/big-brother-gets-a-makeover.html

Bruns, A. (2008). *Blogs, Wikipedia, Second Life, and Beyond: From Production to Produsage*. New York: Peter Lang.

Bruns, A. (2016). Big Data Analysis. In T. Witschge, C. W. Anderson, D. Domingo, & A. Hermida, *The SAGE Handbook of Digital Journalism* (pp. 509–527). London: SAGE Publications. https://doi.org/10.4135/9781473957909.n34

Bruns, A. (2018, April 25). Facebook Shuts the Gate after the Horse Has Bolted, and Hurts Real Research in the Process. Retrieved 2 January 2019, from https://medium.com/@Snurb/facebook-research-data-18662cf2cacb

Brunton, F., & Coleman, G. (2014). Closer to the Metal. In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds.), *Media Technologies: Essays on Communication, Materiality, and Society* (pp. 77–97). Cambridge, MA: MIT

Press. Retrieved from http://mitpress.universitypressscholarship.com/view/10.7551/mitpress/9780262525374.001.0001/upso-9780262525374-chapter-4

Brunton, F., & Nissenbaum, H. (2013). Political and Ethical Perspectives on Data Obfuscation. In M. Hildebrandt & K. de Vries (Eds.), *Privacy, Due Process and the Computational Turn* (pp. 164–188). New York: Routledge.

Bucher, T. (2012a). A Technicity of Attention: How Software 'Makes Sense'. *Culture Machine*, *13*(0). Retrieved from http://www.culturemachine.net/index.php/cm/article/view/470

Bucher, T. (2012b). *Programmed Sociality: A Software Studies Perspective on Social Networking Sites*. University of Oslo, Oslo.

Bucher, T. (2013). Objects of Intense Feeling: The Case of the Twitter API. *Computational Culture*. Retrieved from http://computationalculture.net/objects-of-intense-feeling-the-case-of-the-twitter-api/

Bucher, T. (2018). *If...Then: Algorithmic Power and Politics*. Oxford: Oxford University Press.

Bucher, T., & Helmond, A. (2018). The Affordances of Social Media Platforms. In J. Burgess, A. Marwick, & T. Poell (Eds.), *The SAGE Handbook of Social Media* (pp. 233–253). London: SAGE Publications. https://doi.org/10.4135/9781473984066.n14

Burt, R. S. (1992). *Structural Holes: Social Structure of Competition*. Cambridge, MA: Harvard University Press.

Cairo, A. (2007). *The Elements of Infographics: What Information Graphics Are and Are Not*.

Callon, M. (2004). Europe Wrestling with Technology. *Economy and Society*, *33*(1), 121–134. https://doi.org/10.1080/0308514042000176766

Callon, M., & Muniesa, F. (2005). Peripheral Vision: Economic Markets as Calculative Collective Devices. *Organization Studies*, *26*(8), 1229–1250. https://doi.org/10.1177/0170840605056393

Cardon, D. (2016). Deconstructing the Algorithm Four Types of Digital Information Calculations. In Robert Seyfert & Jonathan Roberge (Eds.), *Algorithmic Cultures Essays on Meaning, Performance and New Technologies*. London: Routledge.

Carlson, M. (2018). Confronting Measurable Journalism. *Digital Journalism*, *6*(4), 406–417. https://doi.org/10.1080/21670811.2018.1445003

Carlson, M., & Lewis, S. C. (Eds.). (2015). *Boundaries of Journalism: Professionalism, Practices and Participation*. London: Routledge.

Castelluccia, C., Grumbach, S., & Olejnik, L. (2013). Data Harvesting 2.0: from the Visible to the Invisible Web, 14.

Chadwick, A. (2013). *The Hybrid Media System: Politics And Power*. Oxford ; New York: Oxford University Press, USA.

Change Advertising. (2016). *The Clickbait Report*. Retrieved from http://changeadvertising.org/the-clickbait-report/

Chilvers, J., & Kearnes, M. (Eds.). (2015). *Remaking Participation: Science, Environment and Emergent Publics* (1 edition). London New York: Routledge.

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. https://doi.org/10.1080/21670811.2014.976400

Couldry, N., & Curran, J. (2003). *Contesting Media Power: Alternative Media in a Networked World*. Lanham, MD: Rowman & Littlefield.

Crawford, K. (2014, July 2). The Test We Can – and Should – Run on Facebook. *The Atlantic*. Retrieved from https://www.theatlantic.com/technology/archive/2014/07/the-test-we-canand-shouldrun-on-facebook/373819/

Culler, J. (2001). *The Pursuit of Signs: Semiotics, Literature, Deconstruction*. London: Routledge.

Dabbish, L., Stuart, C., Tsay, J., & Herbsleb, J. (2012). Social Coding in Github: Transparency and Collaboration in an Open Software Repository (p. 1277). ACM Press. https://doi.org/10.1145/2145204.2145396

Dardenne, R. (2005). Journalism. In D. Herman (Ed.), *Routledge Encyclopedia of Narrative Theory*. London: Routledge.

Darnton, R. (1975). Writing News and Telling Stories. *Daedalus*, *104*(2), 175–194. Retrieved from http://www.jstor.org/stable/20024337

Davies, J. (2016, October 17). 'We've Created a Monster': Publishers Vent Ad

Tech Frustrations. Retrieved 28 May 2018, from https://digiday.com/uk/weve-created-monster-publishers-vent-ad-tech-frustrations/

de Maeyer, J. (2016). Adopting a 'Material Sensibility' in Journalism Studies. In T. Witschge, C. W. Anderson, D. Domingo, & A. Hermida (Eds.), *The SAGE Handbook of Digital Journalism* (pp. 460–476). London: SAGE Publications. https://doi.org/10.4135/9781473957909

de Maeyer, J., & le Cam, F. (2015). The Material Traces of Journalism. *Digital Journalism*, *3*(1), 85–100. https://doi.org/10.1080/21670811.2014.928021

desilva + phillips, LLC. (2008). *Online Ad Networks: Monetizing the Long Tail*. New York. Retrieved from https://dp.oaklins.com/wp-content/uploads/2016/11/DeSilvaPhillips-Ad-Networks-White-Paper-3-08.pdf

Deuze, M. (2012). *Media Life*. Cambridge, UK: Polity Press.

Deuze, M., & Witschge, T. (2018). Beyond Journalism: Theorizing the Transformation of Journalism. *Journalism*, *19*(2), 165–181. https://doi.org/10.1177/1464884916688550

Deville, J., & Velden, L. van der. (2016). Seeing the Invisible Algorithm: The Practical Politics of Tracking the Credit Trackers. In L. Amoore & V. Piotukh (Eds.), *Algorithmic Life: Calculative Devices in the Age of Big Data* (p. 266). London: Routledge.

Dewey, C. (2016, November 17). Facebook Fake-News Writer: 'i Think Donald Trump Is in the White House Because of Me'. *The Washington Post*. Retrieved from https://www.washingtonpost.com/news/the-intersect/wp/2016/11/17/facebook-fake-news-writer-i-think-donald-trump-is-in-the-white-house-because-of-me/?utm_term=.3feaf3edebe2

Diakopoulos, N. (2013). *Algorithmic Accountability Reporting: On the Investigation of Black Boxes*. New York: Tow Center for Digital Journalism. Retrieved from https://doi.org/10.7916/D8ZK5TW2

Diakopoulos, N. (2015). Algorithmic Accountability: Journalistic Investigation of Computational Power Structures. *Digital Journalism*, *3*(3), 398–415. https://doi.org/10.1080/21670811.2014.976411

Digital Methods Initiative. (n.d.-a). *DMI Tools Pertaining to Github*. Retrieved

from https://wiki.digitalmethods.net/Dmi/ToolDatabase?cat=DeviceCentric&subcat=Github

Digital Methods Initiative. (n.d.-b). *Tracker Tracker*. Retrieved from https://tools.digitalmethods.net/beta/trackerTracker/

Domingo, D. (2006). *Inventing Online Journalism: Development of the Internet as a News Medium in Four Catalian Online Newspapers*. Autonomous University of Barcelona, Barcelona. Retrieved from https://ddd.uab.cat/pub/tesis/2006/tdx-1219106-153347/dd1de1.pdf

Domingo, D., Masip, P., & Costera Meijer, I. (2015). Tracing Digital News Networks. *Digital Journalism*, *3*(1), 53–67. https://doi.org/10.1080/21670811.2014.927996

Domingo, D., & Wiard, V. (2016). News Networks. In *The SAGE Handbook of Digital Journalism* (pp. 397–409). London: SAGE Publications. https://doi.org/10.4135/9781473957909

Englehardt, S., & Narayanan, A. (2016). Online Tracking: A 1-Million-Site Measurement and Analysis. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security* (pp. 1388–1401). New York, NY, USA: ACM. https://doi.org/10.1145/2976749.2978313

Espeland, W. (2015). Narrating Numbers. In R. Rottenburg, S. E. Merry, S.-J. Park, & J. Mugler (Eds.), *The World of Indicators: The Making of Governmental Knowledge through Quantification* (pp. 56–75). Cambridge, UK: Cambridge University Press.

Espeland, W. N., & Sauder, M. (2007). Rankings and Reactivity: How Public Measures Recreate Social Worlds. *American Journal of Sociology*, *113*(1), 1–40.

Ettema, J., & Glasser, T. (1998). *Custodians of Conscience: Investigative Journalism and Public Virtue*. New York: Columbia University Press.

Ettema, J. S., & Glasser, T. L. (1988). Narrative Form and Moral Force: The Realization of Innocence and Guilt Through Investigative Journalism. *Journal of Communication*, *38*(3), 8–26. https://doi.org/10.1111/j.1460-2466.1988.tb02057.x

Evans, D. S. (2008). The Economics of the Online Advertising Industry. *Review of Network Economics*, *7*(3), 359–391. Retrieved from

https://papers.ssrn.com/abstract=1086473

Evans, D. S., Hagiu, A., & Schmalensee, R. (2006). *Invisible Engines: How Software Platforms Drive Innovation and Transform Industries*. Cambridge, MA: MIT Press.

Felt, U., Fouché, R., Miller, C. A., Smith-Doerr, L., & Law, J. (2017). *The Handbook of Science and Technology Studies* (Fourth Edition). Cambridge, MA: MIT Press.

Fenton, N., & Barassi, V. (2011). Alternative Media and Social Networking Sites: The Politics of Individuation and Political Participation. *The Communication Review*, *14*(3), 179–196. https://doi.org/10.1080/10714421.2011.597245

Finley, K. (2018, June 30). Github Developers Are Giving Microsoft a Chance. *Wired*. Retrieved from https://www.wired.com/story/github-developers-are-giving-microsoft-a-chance/

Flew, T., Spurgeon, C., Daniel, A., & Swift, A. (2012). The Promise of Computational Journalism. *Journalism Practice*, *6*(2), 157–171. https://doi.org/10.1080/17512786.2011.616655

Ford, H. (2015). Infoboxes and Cleanup Tags: Artifacts of Wikipedia Newsmaking. *Journalism*, *16*(1), 79–98. https://doi.org/10.1177/1464884914545739

Fortunati, L., Taipale, S., & Farinosi, M. (2015). Print and Online Newspapers as Material Artefacts. *Journalism*, *16*(6), 830–846. https://doi.org/10.1177/1464884914545439

Foucault, M. (1980). *Power/Knowledge: Selected Interviews and Other Writings, 1972-1977*. New York: Pantheon Books.

Freeman, L. C. (1982). Centered Graphs and the Structure of Ego Networks. *Mathematical Social Sciences*, *3*(3), 291–304. https://doi.org/10.1016/0165-4896(82)90076-2

Fuller, M. (Ed.). (2008). *Software Studies: A Lexicon*. Cambridge, MA: MIT Press.

Fuller, M., Goffey, A., Mackenzie, A., Mills, R., & Sharples, S. (2017). Big Diff, Granularity, Incoherence, and Production in the Github Software Repository. In *How to be a Geek: Essays on the Culture of Software*. Cambridge, UK: Polity Press.

Gane, N. (2014). The Emergence of Neoliberalism: Thinking Through and

Beyond Michel Foucault's Lectures on Biopolitics. *Theory, Culture & Society*, *31*(4), 3–27. https://doi.org/10.1177/0263276413506944

Gannes, L. (2012, July 9). GitHub Valued at $750M With First Outside Funding Ever. *All Things D*. Retrieved from http://allthingsd.com/20120709/github-valued-at-750m-with-first-outside-funding-ever/

Gehl, R. W. (2011). The Archive and the Processor: The Internal Logic of Web 2.0. *New Media & Society*, *13*(8), 1228–1244. https://doi.org/10.1177/1461444811401735

Gehl, R. W. (2014). *Reverse Engineering Social Media: Software, Culture, and Political Economy in New Media Capitalism*. Philadelphia, PA: Temple University Press.

Genette, G. (1983). *Narrative Discourse: An Essay in Method*. Ithaca: Cornell University Press.

Gerlitz, C. (2016). What Counts? Reflections on the Multivalence of Social Media Data. *Digital Culture & Society*, *2*(2), 19–38. https://doi.org/10.14361/dcs-2016-0203

Gerlitz, C., & Helmond, A. (2013). The Like Economy: Social Buttons and the Data-Intensive Web. *New Media & Society*, *15*(8), 1348–1365. https://doi.org/10.1177/1461444812472322

Gerlitz, C., & Lury, C. (2014). Social Media and Self-Evaluating Assemblages: On Numbers, Orderings and Values. *Distinktion: Journal of Social Theory*, *15*(2), 174–188. https://doi.org/10.1080/1600910X.2014.920267

Gerlitz, C., & Rieder, B. (2018). Tweets Are Not Created Equal: Investigating Twitter's Client Ecosystem. *International Journal of Communication*, *12*, 528–547.

Ghoshal, A. (2016, September 15). Microsoft Has More Open Source Contributors Than Google. Retrieved 30 March 2018, from https://thenextweb.com/microsoft/2016/09/15/in-your-face-google/

Gillespie, T. (2010). The Politics of 'Platforms'. *New Media & Society*, *12*(3), 347–364. https://doi.org/10.1177/1461444809342738

Gillespie, T., Boczkowski, P. J., & Foot, K. A. (2014a). Introduction. In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds.), *Media Technologies: Essays on Communication, Materiality, and Society* (pp. 1–17). Cambridge,

MA: MIT Press. Retrieved from http://www.universitypressscholarship.com/view/10.7551/mitpress/9780262525374.001.0001/upso-9780262525374-chapter-1

Gillespie, T., Boczkowski, P. J., & Foot, K. A. (2014b). *Media Technologies: Essays on Communication, Materiality, and Society*. Cambridge, MA: MIT Press.

GitHub. (2017). The State of the Octoverse 2017. Retrieved 28 March 2018, from https://octoverse.github.com/

GitHub. (n.d.-a). Build Software Better, Together. Retrieved 24 January 2019, from https://github.com/about

GitHub. (n.d.-b). Events - GitHub Developer Guide. Retrieved 20 April 2018, from https://developer.github.com/v3/activity/events/#list-repository-events

GitHub. (n.d.-c). GitHub Help - About GitHub's Use of Your Data. Retrieved 13 April 2018, from https://help.github.com/articles/about-github-s-use-of-your-data/#

GitHub. (n.d.-d). GitHub Help - About Stars. Retrieved 17 April 2018, from https://help.github.com/articles/about-stars/

GitHub. (n.d.-e). GitHub Help - Be Social. Retrieved 19 April 2018, from https://help.github.com/articles/be-social/

GitHub. (n.d.-f). GitHub Help - Fork A Repo. Retrieved 23 April 2018, from https://help.github.com/articles/fork-a-repo/

GitHub. (n.d.-g). GitHub Help - GitHub Privacy Statement. Retrieved 13 April 2018, from https://help.github.com/articles/github-privacy-statement/

GitHub. (n.d.-h). GitHub Help - Terms of Service - License Grant to Other Users. Retrieved 13 April 2018, from https://help.github.com/articles/github-terms-of-service/#d-user-generated-content

GitHub. (n.d.-i). GitHub Terms of Service - API Terms. Retrieved 24 April 2018, from https://help.github.com/articles/github-terms-of-service/#h-api-terms

GitHub. (n.d.-j). The Largest Open Source Community in the World. Retrieved 23 April 2018, from https://github.com/open-source

GitLab. (2018, June 3). We're seeing 10x the normal daily amount of repositories #movingtogitlab https://www.dropbox.com/s/uzg9vc5oljr8lin/Screenshot%202018-06-03%2015.52.52.png?dl=0 … We're scaling our fleet to try to stay up. Follow the progress on https://monitor.gitlab.net/dashboard/db/github-importer?orgId=1 … and @movingtogitlab. *@gitlab*. Tweet. Retrieved 15 January 2019, from https://twitter.com/gitlab/status/1003409836170547200

Gousios, G., Vasilescu, B., Serebrenik, A., & Zaidman, A. (2014). Lean Ghtorrent: Github Data on Demand (pp. 384–387). ACM Press. https://doi.org/10.1145/2597073.2597126

Governor, J. (2017, October 26). So Where All the GitHub Link Farms at? Astro-Turfing in Software Development. Retrieved 28 March 2018, from https://www.enterpriseirregulars.com/119696/github-link-farms-astro-turfing-software-development/

Graefe, A. (2016). *Guide to Automated Journalism*. New York: Tow Center for Digital Journalism. Retrieved from https://towcenter.gitbooks.io/guide-to-automated-journalism/content/

Granovetter, M. S. (1973). The Strength of Weak Ties. *American Journal of Sociology*, *78*(6), 1360–1380. Retrieved from http://www.jstor.org/stable/2776392

Graves, L. (2007). The Affordances of Blogging: A Case Study in Culture and Technological Effects. *Journal of Communication Inquiry*, *31*(4), 331–346. https://doi.org/10.1177/0196859907305446

Gray, J., & Bounegru, L. (Eds.). (forthcoming). *The Data Journalism Handbook: Towards A Critical Data Practice* (2nd edition). Amsterdam: Amsterdam University Press.

Gray, J., Bounegru, L., & Chambers, L. (Eds.). (2012). *The Data Journalism Handbook*. Sebastopol, CA: O'Reilly Media.

Gray, J., Bounegru, L., Milan, S., & Ciuccarelli, P. (2016). Ways of Seeing Data: Towards a Critical Literacy for Data Visualizations as Research Objects and Research Devices. In S. Kubitschko & A. Kaun (Eds.), *Innovative*

Methods in Media and Communication Research (pp. 290–325). London: Palgrave Macmillan.

Gray, J., Gerlitz, C., & Bounegru, L. (2018). Data Infrastructure Literacy. *Big Data & Society*, *5*(2), 1–13. https://doi.org/10.1177/2053951718786316

Griffith, E. (2014, August 18). Taboola and Outbrain Are Battling a Bad Reputation (and Each Other). *Fortune*. Retrieved from http://fortune.com/2014/08/18/taboola-outbrain-battle-bad-reputation-each-other/

Guzman, E., Azócar, D., & Li, Y. (2014). Sentiment Analysis of Commit Comments in GitHub: An Empirical Study. In *Proceedings of the 11th Working Conference on Mining Software Repositories* (pp. 352–355). New York, NY, USA: ACM. https://doi.org/10.1145/2597073.2597118

Hamilton, J. T., & Turner, F. (2009). Accountability Through Algorithm: Developing the Field of Computational Journalism. Presented at the Developing the Field of Computational Journalism, Center For Advanced Study in the Behavioral Sciences.

Haraway, D. J. (2016). *Staying with the Trouble: Making Kin in the Chthulucene*. Durham and London: Duke University Press. Retrieved from www.dukeupress.edu/staying-with-the-trouble

Helmond, A. (2015a). The Platformization of the Web: Making Web Data Platform Ready. *Social Media + Society*, *1*(2), 2056305115603080. https://doi.org/10.1177/2056305115603080

Helmond, A. (2015b). *Web As Platform: Data Flows in Social Media*. University of Amsterdam, Amsterdam.

Helmond, A. (2017). Historical Website Ecology: Analyzing Past States of the Web Using Archived Source Code. In N. Brügger (Ed.), *Web 25: Histories from the First 25 Years of the World Wide Web*. New York: Peter Lang.

Helmond, A., Nieborg, D. B., & van der Vlist, F. N. (2017). The Political Economy of Social Data: A Historical Analysis of Platform-Industry Partnerships (pp. 1–5). ACM Press. https://doi.org/10.1145/3097286.3097324

Hiippala, T. (2014). Multimodal Genre Analysis. In S. Norris & C. D. Maier

(Eds.), *Interactions, Images and Texts: A Reader in Multimodality* (pp. 111–123). Berlin: De Gruyter Mouton.

Interactive Advertising Bureau (IAB). (2014). *How an Ad is Served with Real Time Bidding (RTB)*. YouTube. Retrieved from https://www.youtube.com/watch?v=-Glgi9RRuJs

Introna, L. D. (2016). The Algorithmic Choreography of the Impressionable Subject. In R. Seyfert & J. Roberge (Eds.), *Algorithmic Cultures: Essays on Meaning, Performance and New Technologies*. London: Routledge.

Jacomy, M., Venturini, T., Heymann, S., & Bastian, M. (2014). ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. *PLOS ONE*, *9*(6), e98679. https://doi.org/10.1371/journal.pone.0098679

Jasanoff, S. (2004). Ordering Knowledge, Ordering Society. In S. Jasanoff (Ed.), *States of Knowledge The Co-Production of Science and the Social Order* (pp. 13–45). London: Routledge. Retrieved from https://www.taylorfrancis.com/books/e/9781134328345/chapters/10.4324%2F9780203413845-7

Jensen, P., Morini, M., Karsai, M., Venturini, T., Vespignani, A., Jacomy, M., … Fleury, E. (2015). Detecting Global Bridges in Networks. *Journal of Complex Networks*, *4*(3), 319–329. https://doi.org/10.1093/comnet/cnv022

Jewitt, C. (2009a). An Introduction to Multimodality. In C. Jewitt (Ed.), *The Routledge Handbook of Multimodal Analysis*. London: Routledge.

Jewitt, C. (Ed.). (2009b). *The Routledge Handbook of Multimodal Analysis*. London: Routledge.

Joe, R. (2015, February 27). Everything You Wanted To Know About LiveRamp, The Data Connecter Everyone Uses. Retrieved 29 May 2018, from https://adexchanger.com/data-exchanges/everything-you-wanted-to-know-about-liveramp-the-data-connecter-everyone-uses/

Jonathan Gray, Bounegru, L., & Venturini, T. (forthcoming). 'Fake News' as Infrastructural Uncanny. *New Media & Society*.

Joseph, S. (2018, May 2). Publishers Say They'll Use GDPR to Shed Ad Tech Vendors. Retrieved 28 May 2018, from https://digiday.com/uk/publishers-say-theyll-use-gdpr-shed-ad-tech-

vendors/

Kalliamvakou, E., Gousios, G., Blincoe, K., Singer, L., German, D. M., & Damian, D. (2014). The Promises and Perils of Mining Github (pp. 92–101). ACM Press. https://doi.org/10.1145/2597073.2597074

Karlsson, M. (2010). Rituals of Transparency: Evaluating online news outlets' uses of transparency rituals in the United States, United Kingdom and Sweden. *Journalism Studies*, *11*(4), 535–545. https://doi.org/10.1080/14616701003638400

Karlsson, M., Bergström, A., Clerwall, C., & Fast, K. (2015). Participatory journalism – the (r)evolution that wasn't. Content and user behavior in Sweden 2007–2013. *Journal of Computer-Mediated Communication*, *20*(3), 295–311. https://doi.org/10.1111/jcc4.12115

Keegan, B. C. (2014, April 7). The Need for Openness in Data Journalism. Retrieved 15 March 2018, from https://www.brianckeegan.com/2014/04/the-need-for-openness-in-data-journalism/

Kelty, C. M. (2008). *Two Bits: The Cultural Significance of Free Software*. Durham, NC: Duke University Press.

Khan, H. (2012). Affiliate Platform VigLink Partners With GeoRiot to Enable Location-Based Referrals. Retrieved 18 June 2018, from https://betakit.com/affiliate-platform-viglink-partners-with-georiot-to-enable-location-based-referrals/

Kitchin, R. (2014). Big Data, New Epistemologies and Paradigm Shifts. *Big Data & Society*, *1*(1), 2053951714528481. https://doi.org/10.1177/2053951714528481

Knight, G., & Dean, T. (1982). Myth and the Structure of News. *Journal of Communication*, *32*(2), 144–161. https://doi.org/10.1111/j.1460-2466.1982.tb00502.x

Kormelink, T. G., & Meijer, I. C. (2015). Truthful or Engaging? *Digital Journalism*, *3*(2), 158–174. https://doi.org/10.1080/21670811.2014.1002514

Krebs, V. (2002). Mapping Networks of Terrorist Cells. *Connections*, *24*(3), 43–52.

Kreiss, D. (2015). Afterword. *Journalism*, *16*(1), 153–156.

https://doi.org/10.1177/1464884914545744

Kress, G., & Leeuwen, T. van. (2001). *Multimodal Discourse: The Modes and Media of Contemporary Communication*. London: Bloomsbury Academic.

Kress, Gunther R. (2009). What Is a Mode? In C. Jewitt (Ed.), *The Routledge Handbook of Multimodal Analysis* (pp. 60–75). London: Routledge.

Langlois, G., & Elmer, G. (2013). The Research Politics of Social Media Platforms. *Culture Machine*, 1–17. Retrieved from http://www.culturemachine.net/index.php/cm/article/viewArticle/505

Langlois, G., Elmer, G., McKelvey, F., & Devereaux, Z. (2009). Networked Publics: The Double Articulation of Code and Politics on Facebook. *Canadian Journal of Communication*, *34*(3), 415–434. https://doi.org/10.22230/cjc.2009v34n3a2114

Langlois, G., McKelvey, F., Elmer, G., & Werbin, K. (2009). FCJ-095 Mapping Commercial Web 2.0 Worlds: Towards a New Critical Ontogenesis. *Fibreculture Journal*, (14). Retrieved from http://fourteen.fibreculturejournal.org/fcj-095-mapping-commercial-web-2-0-worlds-towards-a-new-critical-ontogenesis/

Latour, B. (1990). Drawing Things Together. In M. Lynch & S. Woolgar (Eds.), *Representation in Scientific Practice* (pp. 19–68). Cambridge, MA: MIT Press.

Latour, B. (2005a). *Critical Distance or Critical Proximity?* Retrieved from http://www.bruno-latour.fr/sites/default/files/P-113-HARAWAY.pdf

Latour, B. (2005b). *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford: Oxford University Press.

Latour, B., & Woolgar, S. (1986). *Laboratory Life: The Construction of Scientific Facts*. Princeton, NJ: Princeton University Press.

Law, J. (2004). *After Method: Mess in Social Science Research*. London: Routledge.

Law, J., & Ruppert, E. (2013). The Social Life of Methods: Devices. *Journal of Cultural Economy*, *6*(3), 229–240. https://doi.org/10.1080/17530350.2013.812042

Law, J., Ruppert, E., & Savage, M. (2011). *The Double Social Life of Methods* (Working Paper) (pp. 1–18). Milton Keynes: Centre for Research on

Socio-Cultural Change (CRESC). Retrieved from http://research.gold.ac.uk/7987/

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., … Alstyne, M. V. (2009). Computational Social Science. *Science*, *323*(5915), 721–723. https://doi.org/10.1126/science.1167742

Lehtiniemi, T. (2016). Reorienting Datafication? New Roles for Users in Online Platform Markets (p. 16). Presented at the Internet, Policy and Politics Conference, University of Oxford, Oxford, UK.

Leonardi, P. M. (2012). Materiality, Sociomateriality, and Socio-Technical Systems: What Do These Terms Mean? How Are They Different? Do We Need Them? In P. M. Leonardi, B. A. Nardi, & J. Kallinikos (Eds.), *Materiality and Organizing* (pp. 24–48). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199664054.003.0002

Les Décodeurs. (2017, January 23). L'annuaire Des Sources Du Décodex: Mode D'Emploi. Retrieved from https://www.lemonde.fr/les-decodeurs/article/2017/01/23/l-annuaire-des-sources-du-decodex-mode-d-emploi_5067719_4355770.html

Lewis, S. C., & Usher, N. (2013). Open Source and Journalism: Toward New Frameworks for Imagining News Innovation. *Media, Culture & Society*, *35*(5), 602–619. https://doi.org/10.1177/0163443713485494

Lewis, S. C., & Westlund, O. (2015). Actors, Actants, Audiences, and Activities in Cross-Media News Work. *Digital Journalism*, *3*(1), 19–37. https://doi.org/10.1080/21670811.2014.927986

Lewis, S. C., & Westlund, O. (2016). Mapping the Human–Machine Divide in Journalism. In T. Witschge, C. W. Anderson, D. Domingo, & A. Hermida (Eds.), *The SAGE Handbook of Digital Journalism* (pp. 341–353). 1 Oliver's Yard, 55 City Road London EC1Y 1SP: SAGE Publications Ltd. https://doi.org/10.4135/9781473957909.n23

Lewis, S. C., Zamith, R., & Hermida, A. (2013). Content Analysis in an Era of Big Data: A Hybrid Approach to Computational and Manual Methods. *Journal of Broadcasting & Electronic Media*, *57*(1), 34–52. https://doi.org/10.1080/08838151.2012.761702

Lezaun, J., Marres, N., & Tironi, M. (2017). Experiments in Participation. In U. Felt, R. Fouché, C. A. Miller, & L. Smith-Doerr (Eds.), *Handbook of*

Science and Technology Studies (4th ed., pp. 195–222). Cambridge MA: MIT Press.

Libert, T. (2015). Exposing the Invisible Web: An Analysis of Third-Party HTTP Requests on 1 Million Websites. *International Journal of Communication*, *9*, 18. Retrieved from http://ijoc.org/index.php/ijoc/article/view/3646

Libert, T. (2018). An Automated Approach to Auditing Disclosure of Third-Party Data Collection in Website Privacy Policies (pp. 207–216). ACM Press. https://doi.org/10.1145/3178876.3186087

Libert, T., & Nielsen, R. K. (2018). *Third-Party Web Content on EU News Sites*. Oxford: Reuters Institute, University of Oxford. Retrieved from https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2018-05/Third-Party%20Web%20Content%20on%20EU%20News%20Sites.pdf

Lichterman, J. (2017, February 9). With $1.1 Million in Funding from Knight, Opennews Is Becoming an Independent Organization. Retrieved 13 April 2018, from http://www.niemanlab.org/2017/02/with-1-1-million-in-funding-from-knight-opennews-is-becoming-an-independent-organization/

Lievrouw, L. A. (2014). Materiality and Media in Communication and Technology Studies: An Unfinished Project. In *Media Technologies: Essays on Communication, Materiality, and Society* (pp. 21–51). Cambridge, MA: MIT Press. Retrieved from http://mitpress.universitypressscholarship.com/view/10.7551/mitpress/9780262525374.001.0001/upso-9780262525374-chapter-2

Lopes, C. V., Maj, P., Martins, P., Saini, V., Yang, D., Zitny, J., … Vitek, J. (2017). DéJàVu: A Map of Code Duplicates on GitHub. *Proc. ACM Program. Lang.*, *1*(OOPSLA), 84:1–84:28. https://doi.org/10.1145/3133908

Lury, C., & Wakeford, N. (Eds.). (2012). *Inventive Methods: The Happening of the Social*. London: Routledge.

Lynch, M. (1985). Discipline and the Material Form of Images: An Analysis of Scientific Visibility, Discipline and the Material Form of Images: An Analysis of Scientific Visibility. *Social Studies of Science*, *15*(1), 37–66.

https://doi.org/10.1177/030631285015001002

Macbeth, S. (2017). *Tracking the Trackers: Analysing the Global Tracking Landscape with Ghostrank* (p. 13). Cliqz.

Mackenzie, A. (2018). 48 Million Configurations and Counting: Platform Numbers and Their Capitalization. *Journal of Cultural Economy*, *11*(1), 36–53. https://doi.org/10.1080/17530350.2017.1393443

Mackenzie, A., Fuller, M., Goffey, A., Mills, R., & Sharples, S. (2014). *Metacommunities of Practice in the Code-Sharing Commons* (Discussion paper). Retrieved from http://metacommunities.github.io/metacommunities/

Mackenzie, A., Mills, R., Sharples, S., Fuller, M., & Goffey, A. (2015). Digital Sociology in the Field of Devices. In L. Hanquinet & M. Savage (Eds.), *Routledge International Handbook of the Sociology of Art and Culture*. London: Routledge. Retrieved from https://www.taylorfrancis.com/books/e/9781135008895/chapters/10.4324%2F9780203740248-20

Manovich, L. (2001). *The Language of New Media*. MIT Press.

Manovich, L. (2008). *Software Takes Command*. London: Bloomsbury.

Marlow, J., Dabbish, L., & Herbsleb, J. (2013). Impression Formation in Online Peer Production: Activity Traces and Personal Profiles in Github. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (pp. 117–128). New York, NY, USA: ACM. https://doi.org/10.1145/2441776.2441792

Marres, N. (2011). The Costs of Public Involvement: Everyday Devices of Carbon Accounting and the Materialization of Participation. *Economy and Society*, *40*(4), 510–533. https://doi.org/10.1080/03085147.2011.602294

Marres, N. (2012a). *Material Participation: Technology, the Environment and Everyday Publics*. Houndmills, Basingstoke: Palgrave Macmillan.

Marres, N. (2012b). The Redistribution of Methods: On Intervention in Digital Social Research, Broadly Conceived. *The Sociological Review*, *60*(S1), 139–165. https://doi.org/10.1111/j.1467-954X.2012.02121.x

Marres, N. (2013). *Who Is Afraid of the Green Cloud? On the Environmental Rendering of Controversy* (Centre for the Study of Invention & Social Process (CSISP) No. 2). London: Goldsmiths, University of London.

Marres, N. (2017a). *Digital Sociology: The Reinvention of Social Research*. Cambridge, UK: Polity Press.

Marres, N. (2017b, March). *The Democracy Test: Street Trials as Experiments in Interpretation*. Presented at the Sonic Acts Festival: The Noise of Being, Amsterdam. Retrieved from https://vimeo.com/210172639

Marres, N., & Gerlitz, C. (2015). Interface Methods: Renegotiating Relations Between Digital Social Research, Sts and Sociology. *The Sociological Review*, *64*(1), 21–46. https://doi.org/10.1111/1467-954X.12314

Marres, N., Guggenheim, M., & Wilkie, A. (Eds.). (2018). *Inventing the Social*. Manchester: Mattering Press.

Marres, N., & Moats, D. (2015). Mapping Controversies with Social Media: The Case for Symmetry. *Social Media + Society*, *1*(2), 2056305115604176. https://doi.org/10.1177/2056305115604176

Marres, N., & Weltevrede, E. (2013). Scraping the Social? Issues in Live Social Research. *Journal of Cultural Economy*, *6*(3), 313–335. https://doi.org/10.1080/17530350.2013.772070

Marvin, G. (2016, April 22). What Is a Data Management Platform? Retrieved 18 June 2018, from https://martechtoday.com/what-is-dmp-martech-landscape-174298

Matheson, D. (2004). Weblogs and the Epistemology of the News: Some Trends in Online Journalism , Weblogs and the Epistemology of the News: Some Trends in Online Journalism. *New Media & Society*, *6*(4), 443–468. https://doi.org/10.1177/146144804044329

Mayer, J. R., & Mitchell, J. C. (2012). Third-Party Web Tracking: Policy and Technology (pp. 413–427). IEEE. https://doi.org/10.1109/SP.2012.47

McChesney, R. W. (2003). The Problem of Journalism: a political economic contribution to an explanation of the crisis in contemporary US journalism. *Journalism Studies*, *4*(3), 299–329. https://doi.org/10.1080/14616700306492

McDonald, N., & Goggins, S. (2013). Performance and Participation in Open Source Software on GitHub. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (pp. 139–144). New York, NY, USA: ACM. https://doi.org/10.1145/2468356.2468382

McNair, B. (2017). *Fake News: Falsehood, Fabrication and Fantasy in Journalism*. London: Routledge.

Meehan, E. R. (1984). Ratings and the Institutional Approach: A Third Answer to the Commodity Question. *Critical Studies in Mass Communication*, *1*(2), 216–225. https://doi.org/10.1080/15295038409360032

Micó, J. L., Masip, P., & Domingo, D. (2013). To Wish Impossible Things*: Convergence as a Process of Diffusion of Innovations in an Actor-Network. *International Communication Gazette*, *75*(1), 118–137. https://doi.org/10.1177/1748048512461765

Mitchelstein, E., & Boczkowski, P. J. (2009). Between Tradition and Change: A Review of Recent Research on Online News Production. *Journalism*, *10*(5), 562–586. https://doi.org/10.1177/1464884909106533

Moats, D. (2016). *Decentring Devices: Developing Quali-Quantitative Techniques for Studying Controversies with Online Platforms*. University of London, London. Retrieved from https://research.gold.ac.uk/17041/1/SOC_thesis_MoatsD_2016.pdf

Moats, D., & Borra, E. (2018). Quali-Quantitative Methods Beyond Networks: Studying Information Diffusion on Twitter with the Modulation Sequencer. *Big Data & Society*, *5*(1), 2053951718772137. https://doi.org/10.1177/2053951718772137

Mol, A. (1999). Ontological Politics. A Word and Some Questions. *The Sociological Review*, *47*(S1), 74–89. https://doi.org/10.1111/j.1467-954X.1999.tb03483.x

Moreno, J. L. (1977). *Who Shall Survive?: Foundations of Sociometry, Group Psychotherapy, and Sociodrama*. Beacon, NY: Beacon House.

Moretti, F. (2011). *Network Theory, Plot Analysis* (Literary Lab No. Pamphlet 2) (p. 42).

Morgan, M. S. (2012). Case Studies: One Observation or Many? Justification or Discovery? *Philosophy of Science*, *79*(5), 667–677. https://doi.org/10.1086/667848

Mosco, V. (2009). *The Political Economy of Communication*. London: SAGE Publications.

Moses, L. (2016, November 28). 'The Underbelly of the Internet': How Content Ad Networks Fund Fake News. Retrieved 31 January 2017,

from http://digiday.com/publishers/underbelly-internet-fake-news-gets-funded/

Muniesa, F., Millo, Y., & Callon, M. (2007). An Introduction to Market Devices. In M. Callon, Y. Millo, & F. Muniesa (Eds.), *Market Devices* (pp. 1–12). Chichester: Wiley-Blackwell. Retrieved from https://www.wiley.com/en-ad/Market+Devices-p-9781405170284

Nagy, P., & Neff, G. (2015). Imagined Affordance: Reconstructing a Keyword for Communication Theory. *Social Media + Society*, *1*(2), 2056305115603385. https://doi.org/10.1177/2056305115603385

Napoli, P. (2011). *Audience Evolution: New Technologies and the Transformation of Media Audiences*. New York: Columbia University Press.

Napoli, P. M. (2003). *Audience Economics: Media Institutions and the Audience Marketplace*. New York: Columbia University Press.

Nechushtai, E. (2018). Could Digital Platforms Capture the Media Through Infrastructure? *Journalism*, *19*(8), 1043–1058. https://doi.org/10.1177/1464884917725163

Neculai, D. (2017, November 2). How to Get up to 3500+ Github Stars in One Week – and Why That's Important. Retrieved 29 March 2018, from https://medium.freecodecamp.org/how-to-get-up-to-3500-github-stars-in-one-week-339102b62a8f

Neff, G. (2015). Learning from Documents: Applying New Theories of Materiality to Journalism. *Journalism*, *16*(1), 74–78. https://doi.org/10.1177/1464884914549294

Neff, G., Jordan, T., McVeigh-Schultz, J., & Gillespie, T. (2012). Affordances, Technical Agency, and the Politics of Technologies of Cultural Production. *Journal of Broadcasting & Electronic Media*, *56*(2), 299–313. https://doi.org/10.1080/08838151.2012.678520

Nickel, B. (n.d.). The Dark Side of Display: How "Fake News" Sites Monetize their Content. Retrieved 29 June 2018, from https://blog.adbeat.com/fake-news/

Nieborg, D. B. (2016). Free-to-Play Games and App Advertising. In J. F. Hamilton, R. Bodle, & E. Korin (Eds.), *Explorations in Critical Studies of Advertising*. London: Routledge.

Nieborg, D. B., & Poell, T. (2018). The Platformization of Cultural

Production: Theorizing the Contingent Cultural Commodity. *New Media & Society*, *20*(11), 4275–4292. https://doi.org/10.1177/1461444818769694

Nielsen, R. K. (2016). The Business of News. In T. Witschge, C. W. Anderson, D. Domingo, & A. Hermida (Eds.), *The SAGE Handbook of Digital Journalism* (pp. 51–67). London: SAGE Publications. https://doi.org/10.4135/9781473957909.n4

Nielsen, R. K. (2018). The Changing Economic Contexts of Journalism. In K. Wahl-Jorgensen & T. Hanitzsch (Eds.), *Handbook of Journalism Studies*. London: Routledge. Retrieved from https://rasmuskleisnielsen.files.wordpress.com/2018/05/nielsen-the-changing-economic-contexts-of-journalism-v2.pdf

Nielsen, R. K., & Ganter, S. A. (2018). Dealing with digital intermediaries: A case study of the relations between publishers and platforms. *New Media & Society*, *20*(4), 1600–1617. https://doi.org/10.1177/1461444817701318

Noack, A. (2009). Modularity Clustering Is Force-Directed Layout. *Physical Review E*, *79*(2). https://doi.org/10.1103/PhysRevE.79.026102

O'Reilly, T. (2007). What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. *International Journal of Digital Economics*, (65), 17–37. Retrieved from https://mpra.ub.uni-muenchen.de/4580/

Papacharissi, Z., & Oliveira, M. de F. (2012). Affective News and Networked Publics: The Rhythms of News Storytelling on #Egypt. *Journal of Communication*, *62*(2), 266–282. https://doi.org/10.1111/j.1460-2466.2012.01630.x

Parikka, J. (2011, May 15). Platform Politics-Conference: Opening Words. Retrieved 27 March 2018, from https://jussiparikka.net/2011/05/15/platform-politics-conference-opening-words/

Pavlik, J. (2000). The Impact of Technology on Journalism. *Journalism Studies*, *1*(2), 229–237. https://doi.org/10.1080/14616700050028226

Pavlik, John V., & Bridges, F. (2013). The Emergence of Augmented Reality (AR) as a Storytelling Medium in Journalism. *Journalism &*

Communication Monographs, *15*(1), 4–59. https://doi.org/10.1177/1522637912470819

Pavlik, John V., & Pavlik, J. O. (2017). Understanding Quality in Digital Storytelling: A Theoretically Based Analysis of the Interactive Documentary. In M. Friedrichsen & Y. Kamalipour (Eds.), *Digital Transformation in Journalism and News Media: Media Management, Media Convergence and Globalization* (pp. 381–396). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-27786-8_28

Pavlik, John Vernon. (2001). *Journalism and New Media*. Columbia University Press.

Pentland, B. T., & Feldman, M. S. (2007). Narrative Networks: Patterns of Technology and Organization. *Organization Science*, *18*(5), 781–795. https://doi.org/10.1287/orsc.1070.0283

Peters, C., & Broersma, M. (2016). *Rethinking Journalism Again: Societal role and public relevance in a digital age*. Routledge.

Peters, C., & Witschge, T. (2015). From Grand Narratives of Democracy to Small Expectations of Participation. *Journalism Practice*, *9*(1), 19–34. https://doi.org/10.1080/17512786.2014.928455

Peters, J. D. (2011, March). *Two Cheers for Technological Determinism*. Presented at the Conference on Media Histories: Epistemology, Materiality, Temporality, New York. Retrieved from https://vimeo.com/25591045

Peters, J. D. (2017). "You Mean My Whole Fallacy Is Wrong": On Technological Determinism. *Representations*, *140*(1), 10–26. https://doi.org/10.1525/rep.2017.140.1.10

Petersen, B. (2017, October 12). Github, the 'Facebook for Programmers,' Has Quietly Built up an Enterprise Business That Accounts for Half Its $200 Million in Sales. *Business Insider*. Retrieved from http://uk.businessinsider.com/githubs-enterprise-service-brings-in-half-its-200-million-in-sales-2017-10

Petre, C. (2015). *The Traffic Factories: Metrics at Chartbeat, Gawker Media, and the New York Times* (A Tow/Knight Report). New York: Tow Center for Digital Journalism. Retrieved from https://towcenter.org/research/traffic-factories/

Pickard, V. (2011). Can Government Support the Press? Historicizing and Internationalizing a Policy Approach to the Journalism Crisis. *The Communication Review*, *14*(2), 73–95. https://doi.org/10.1080/10714421.2011.573430

Pickard, V. (2014, July 1). Wither(ing) Journalism? *Public Books*. Retrieved from http://www.publicbooks.org/withering-journalism/

Pickard, V. (2017). Rediscovering the News: Journalism Studies' Three Blind Spots. In P. J. Boczkowski & C. W. Anderson (Eds.), *Remaking the News: Essays on the Future of Journalism Scholarship in the Digital Age*. Cambridge, MA: MIT Press.

Pidgeon, D. (2016, October 4). Where Did the Money Go? Guardian Buys Its Own Ad Inventory. Retrieved 27 June 2018, from https://mediatel.co.uk/newsline/2016/10/04/where-did-the-money-go-guardian-buys-its-own-ad-inventory

Plantin, J.-C., Lagoze, C., Edwards, P. N., & Sandvig, C. (2016). Infrastructure Studies Meet Platform Studies in the Age of Google and Facebook. *New Media & Society*, *20*(1), 293–310. https://doi.org/10.1177/1461444816661553

Plesner, U. (2009). An Actor-Network Perspective on Changing Work Practices: Communication Technologies as Actants in Newswork. *Journalism*, *10*(5), 604–626. https://doi.org/10.1177/1464884909106535

Poell, T., & Borra, E. (2012). Twitter, YouTube, and Flickr as platforms of alternative journalism: The social media account of the 2010 Toronto G20 protests. *Journalism*, *13*(6), 695–713. https://doi.org/10.1177/1464884911431533

Power, M. (1994). *The Audit Explosion*. London: Demos.

Power, M. (1999). *The Audit Society: Rituals of Verification*. Oxford: Oxford University Press.

Power, M. (2000). The Audit Society - Second Thoughts. *International Journal of Auditing*, *4*(1), 111–119. https://doi.org/10.1111/1099-1123.00306

Primo, A., & Zago, G. (2015). Who And What Do Journalism? *Digital Journalism*, *3*(1), 38–52. https://doi.org/10.1080/21670811.2014.927987

Puschmann, C., & Burgess, J. (2013). The Politics of Twitter Data. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2206225

Ragin, C. C. (1992). 'Casing' and the Process of Social Inquiry. In H. S. Becker & C. C. Ragin (Eds.), *What is a case? Exploring the foundations of social inquiry* (pp. 217–226). Cambridge: Cambridge University Press. Retrieved from https://trove.nla.gov.au/version/45429805

Ratliff, J. D., & Rubinfeld, D. L. (2010). Online Advertising: Defining Relevant Markets. *Journal of Competition Law & Economics*, *6*(3), 653–686.

Ray, B., Posnett, D., Filkov, V., & Devanbu, P. (2014). A Large Scale Study of Programming Languages and Code Quality in Github. In *Proceedings of the 22Nd ACM SIGSOFT International Symposium on Foundations of Software Engineering* (pp. 155–165). New York, NY, USA: ACM. https://doi.org/10.1145/2635868.2635922

Richard Serra. (1973). *Television Delivers People*. Retrieved from https://www.youtube.com/watch?v=LvZYwaQlJsg

Richtel, M. (2013, April 27). How Big Data Is Playing Recruiter for Specialized Workers. *The New York Times*. Retrieved from https://www.nytimes.com/2013/04/28/technology/how-big-data-is-playing-recruiter-for-specialized-workers.html

Rieder, B. (2012). What is in PageRank? A Historical and Conceptual Investigation of a Recursive Status Index. *Computational Culture*. Retrieved from http://computationalculture.net/what_is_in_pagerank/

Rieder, B. (2013, July 10). Scrutinizing a Network of Likes on Facebook (and Some Thoughts on Network Analysis and Visualization). Retrieved 24 January 2019, from http://thepoliticsofsystems.net/2013/07/scrutinizing-a-network-of-likes-on-facebook-and-some-thoughts-on-network-analysis-and-visualization/

Rieder, B. (2017). Beyond Surveillance: How Do Markets and Algorithms 'Think'? *Le Foucaldien*, *3*(1), 1–20. https://doi.org/10.16995/lefou.30

Rieder, B., & Sire, G. (2014). Conflicts of Interest and Incentives to Bias: A Microeconomic Critique of Google's Tangled Position on the Web. *New Media & Society*, *16*(2), 195–211.

https://doi.org/10.1177/1461444813481195

Rochet, J.-C., & Tirole, J. (2003). Platform Competition in Two-Sided Markets. *Journal of the European Economic Association*, *1*(4), 990–1029. https://doi.org/10.1162/154247603322493212

Rodgers, S. (2015). Foreign Objects? Web Content Management Systems, Journalistic Cultures and the Ontology of Software. *Journalism*, *16*(1), 10–26. https://doi.org/10.1177/1464884914545729

Roesner, F., Kohno, T., & Wetherall, D. (2012). Detecting and Defending Against Third-Party Tracking on the Web. In *Proceedings of the 9th USENIX Conference on Networked Systems Design and Implementation* (pp. 12–12). Berkeley, CA, USA: USENIX Association. Retrieved from http://dl.acm.org/citation.cfm?id=2228298.2228315

Rogers, R. (2002). Operating Issue Networks on the Web. *Science as Culture*, *11*(2), 191–214.

Rogers, R. (2004). *Information Politics on the Web*. Cambridge, MA: MIT Press.

Rogers, R. (2013). *Digital Methods*. Cambridge, MA: MIT Press.

Rogers, R. (2014). Political Research in the Digital Age. *International Public Policy Review*, *8*(1), 73–87. Retrieved from https://www.ucl.ac.uk/ippr/journal/vol8no1

Rogers, R. (2017). Digital Methods for Cross Platform Analysis. In J. Burgess, A. Marwick, & T. Poell (Eds.), *The SAGE Handbook of Social Media*. London: SAGE Publications.

Rogers, R. (2018). Social Media Research After the Fake News Debacle. *Partecipazione e Conflitto*, *11*(2), 557-570–570. https://doi.org/10.1285/i20356609v11i2p557

Rogers, R., & Marres, N. (2000). Landscaping Climate Change: A Mapping Technique for Understanding Science and Technology Debates on the World Wide Web. *Public Understanding of Science*, *9*(2), 141–163. https://doi.org/10.1088/0963-6625/9/2/304

Rohan, J. (2013, August 13). Explore What Is Trending on Github. Retrieved 17 April 2018, from https://blog.github.com/2013-08-13-explore-what-is-trending-on-github/

Ruppert, E., Law, J., & Savage, M. (2013). Reassembling Social Science Methods: The Challenge of Digital Devices. *Theory, Culture & Society*,

30(4), 22–46. https://doi.org/10.1177/0263276413484941

Russell, M. A. (2013). *Mining the Social Web: Data Mining Facebook, Twitter, LinkedIn, Google+, GitHub, and More*. Sebastopol, CA: O'Reilly Media.

Ryan, M.-L. (2005). Narrative and Digitality: Learning to Think With the Medium. In J. Phelan & P. J. Rabinowitz (Eds.), *A Companion To Narrative Theory* (pp. 515–528). Oxford: Wiley-Blackwell.

Ryan, M.-L., & Ryan, M.-L. (Eds.). (2004). Introduction. In *Narrative across Media: The Languages of Storytelling*. Lincoln, NE: University of Nebraska Press.

Savage, M., & Burrows, R. (2007). The Coming Crisis of Empirical Sociology. *Sociology*, *41*(5), 885–899. https://doi.org/10.1177/0038038507080443

Savage, M., Ruppert, E. S., & Law, J. (2010). *Digital Devices: Nine Theses*. London: CRESC.

Schmitz Weiss, A., & Domingo, D. (2010). Innovation Processes in Online Newsrooms as Actor-Networks and Communities of Practice. *New Media & Society*, *12*(7), 1156–1171. https://doi.org/10.1177/1461444809360400

Schroeder, R. (2004). Interactive Info Graphics in Europe – Added Value to Online Mass Media: A Preliminary Survey. *Journalism Studies*, *5*(4), 563–570. https://doi.org/10.1080/14616700412331296473

Schudson, M. (1982). The Politics of Narrative Form: The Emergence of News Conventions in Print and Television. *Daedalus*, *111*(4), 97–112. Retrieved from http://www.jstor.org/stable/20024819

Scudder, D., & Wiener, D. (2012). A Primer on Data Onboarding. Retrieved 18 June 2018, from https://liveramp.com/blog/a-primer-on-data-onboarding/

Segel, E., & Heer, J. (2010). Narrative Visualization: Telling Stories with Data. *IEEE Transactions on Visualization and Computer Graphics*, *16*(6), 1139–1148. https://doi.org/10.1109/TVCG.2010.179

Siles, I., & Boczkowski, P. J. (2012). Making sense of the newspaper crisis: A critical assessment of existing research and an agenda for future work. *New Media & Society*, *14*(8), 1375–1394. https://doi.org/10.1177/1461444812455148

Silverman, C. (2016, November 16). This Analysis Shows How Viral Fake

Election News Stories Outperformed Real News On Facebook. Retrieved 5 July 2018, from https://www.buzzfeed.com/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook

Silverman, C., & Alexander, L. (2016, November 3). How Teens In The Balkans Are Duping Trump Supporters With Fake News. *BuzzFeed*. Retrieved from https://www.buzzfeed.com/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trump-misinfo

Silverman, C., Singer-Vine, J., & Vo, L. T. (2017, April 4). Fake News, Real Ads: Fake News Publishers Are Still Earning Money From Big Ad Networks. Retrieved 5 July 2018, from https://www.buzzfeed.com/craigsilverman/fake-news-real-ads

Singer, J. B., Domingo, D., Heinonen, A., Hermida, A., Paulussen, S., Quandt, T., … Vujnovic, M. (2011). *Participatory Journalism: Guarding Open Gates at Online Newspapers*. John Wiley & Sons.

Sismondo, S. (2008). Science and Technology Studies and an Engaged Program. In Edward J. Hackett, Olga Amsterdamska, Michael Lynch, & Judy Wajcman (Eds.), *The Handbook of Science and Technology Studies* (Third Edition, pp. 13–32). Cambridge, MA: MIT Press.

Sismondo, S. (2017). Post-Truth? *Social Studies of Science*, *47*(1), 3–6. https://doi.org/10.1177/0306312717692076

Smythe, D. W. (1977). Communications: Blindspot of Western Marxism. *Canadian Journal of Political and Social Theory*, *1*(3), 1–27. Retrieved from https://journals.uvic.ca/index.php/ctheory/article/view/13715

Soltani, A., Canty, S., Mayo, Q., Thomas, L., & Hoofnagle, C. J. (2009). *Flash Cookies and Privacy* (SSRN Scholarly Paper No. ID 1446862). Rochester, NY: Social Science Research Network. Retrieved from https://papers.ssrn.com/abstract=1446862

Source. (n.d.). About Source - Source: An OpenNews project. Retrieved 25 April 2018, from https://source.opennews.org/about/

Srnicek, N. (2016). *Platform Capitalism*. Cambridge, UK: Polity Press.

Star, S. L. (1990). Power, Technology and the Phenomenology of Conventions: On Being Allergic to Onions. *The Sociological Review*, *38*(S1), 26–56. https://doi.org/10.1111/j.1467-954X.1990.tb03347.x

Stark, J. A., & Diakopoulos, N. (2016). Towards Editorial Transparency in Computational Journalism (p. 5). Presented at the Computation + Journalism Symposium, Palo Alto, California, USA.

Starr, P. (2012). An Unexpected Crisis: The News Media in Postindustrial Democracies. *The International Journal of Press/Politics*, *17*(2), 234–242. https://doi.org/10.1177/1940161211434422

Steensen, S. (2011). Online Journalism and the Promises of New Technology. *Journalism Studies*, *12*(3), 311–327. https://doi.org/10.1080/1461670X.2010.501151

Steensen, S., & Ahva, L. (2015). Theories of Journalism in a Digital Age. *Journalism Practice*, *9*(1), 1–18. https://doi.org/10.1080/17512786.2014.928454

Strathern, M. (Ed.). (2000). *Audit Cultures: Anthropological Studies in Accountability, Ethics and the Academy*. London: Routledge.

Stroud, N. J., Scacco, J. M., & Curry, A. L. (2016). The Presence and Use of Interactive Features on News Websites. *Digital Journalism*, *4*(3), 339–358. https://doi.org/10.1080/21670811.2015.1042982

Subramanian, S. (2017, February 15). Inside the Macedonian Fake-News Complex. *Wired*. Retrieved from https://www.wired.com/2017/02/veles-macedonia-fake-news/

Suchman, L. (2014). Mediations and Their Others. In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds.), *Media Technologies: Essays on Communication, Materiality, and Society* (pp. 129–138). Cambridge, MA: The MIT Press. https://doi.org/10.7551/mitpress/9780262525374.003.0007

Sudhahar, S., Fazio, G. D., Franzosi, R., & Cristianini, N. (2015). Network Analysis of Narrative Content in Large Corpora. *Natural Language Engineering*, *21*(1), 81–112. https://doi.org/10.1017/S1351324913000247

Sullivan, B. (2016, June 1). GitHub Is Relying On Developers To Crack The Business Market. *Silicon UK*. Retrieved from https://www.silicon.co.uk/software/open-source/github-enterprise-coding-open-source-193037

Suri, A. (2017, January 27). How We Got Our 2-Year-Old Repo Trending on

Github in Just 48 Hours. Retrieved 29 March 2018, from https://medium.freecodecamp.org/how-we-got-a-2-year-old-repo-trending-on-github-in-just-48-hours-12151039d78b

Suslik Spritzer, A., Boy, J., Dragicevic, P., Fekete, J.-D., Freitas, D. S., & Maria, C. (2015). Towards a Smooth Design Process for Static Communicative Node-Link Diagrams. https://doi.org/10.1111/cgf.12658

Tandoc, E. C. J., Jenkins, J., & Craft, S. (2018). Fake News as a Critical Incident in Journalism. *Journalism Practice*, *0*(0), 1–17. https://doi.org/10.1080/17512786.2018.1562958

Tandoc Jr., E. C., Lim, Z. W., & Ling, R. (2018). Defining "Fake News". *Digital Journalism*, *6*(2), 137–153. https://doi.org/10.1080/21670811.2017.1360143

Thorsen, E., & Jackson, D. (2018). Seven Characteristics Defining Online News Formats. *Digital Journalism*, *6*(7), 847–868. https://doi.org/10.1080/21670811.2018.1468722

Thrift, N. J. (2005). *Knowing Capitalism*. London: SAGE Publications.

Thung, F., Bissyande, T. F., Lo, D., & Jiang, L. (2013). Network Structure of Social Coding in GitHub (pp. 323–326). IEEE. https://doi.org/10.1109/CSMR.2013.41

Tsay, J., Dabbish, L., & Herbsleb, J. (2014). Influence of Social and Technical Factors for Evaluating Contribution in Github (pp. 356–366). ACM Press. https://doi.org/10.1145/2568225.2568315

Tuchman, G. (1976). Telling Stories. *Journal of Communication*, *26*(4), 93–97. https://doi.org/10.1111/j.1460-2466.1976.tb01942.x

Tufekci, Z. (2014). Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls. *ArXiv:1403.7400 [Physics]*. Retrieved from http://arxiv.org/abs/1403.7400

Tufekci, Z. (2017, October 4). YouTube is the most overlooked story of 2016. It's search and recommender algorithms are misinformation engines. https://twitter.com/Walldo/status/915422157433298945 …. *@zeynep*. Tweet. Retrieved 2 January 2019, from https://twitter.com/zeynep/status/915608049141915648

Turow, J. (2011). *The Daily You*. New Haven, CT: Yale University Press.

Usher, N. (2016). *Interactive Journalism: Hackers, Data, and Code*. Urbana, IL: University of Illinois Press.

Usher, N. (2018). Re-Thinking Trust in the News. *Journalism Studies*, *19*(4), 564–578. https://doi.org/10.1080/1461670X.2017.1375391

Utt, S. H., & Pasternak, S. (2000). Update on Infographics in American Newspapers. *Newspaper Research Journal*, *21*(2), 55–66. https://doi.org/10.1177/073953290002100205

van Couvering, E. (2017). The Political Economy of New Media Revisited. In *Proceedings of the 50th Hawaii International Conference on System Sciences*. Hawaii.

van der Velden, L. (2014). The Third Party Diary: Tracking the Trackers on Dutch Governmental Websites. *NECSUS. European Journal of Media Studies*, *3*(1), 195–217. https://doi.org/10.5117/NECSUS2014.1.VELD

van der Velden, L. (2018). *Surveillance as Public Matter Revisiting Surveillance Through Devices and Leaks*. Universtity of Amsterdam, Amsterdam.

van Dijck, Jose. (2013). *The Culture of Connectivity: A Critical History of Social Media*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199970773.001.0001

van Dijck, José, Poell, T., & de Waal, M. (2018). *The Platform Society: Public Values in a Connective World*. Oxford University Press.

Vasilescu, B., Posnett, D., Ray, B., van den Brand, M. G. J., Serebrenik, A., Devanbu, P., & Filkov, V. (2015). Gender and Tenure Diversity in GitHub Teams. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 3789–3798). New York, NY, USA: ACM. https://doi.org/10.1145/2702123.2702549

Vasilescu, B., Yu, Y., Wang, H., Devanbu, P., & Filkov, V. (2015). Quality and productivity outcomes relating to continuous integration in GitHub (pp. 805–816). ACM Press. https://doi.org/10.1145/2786805.2786850

Venturini, T. (2012). Building on faults: How to represent controversies with digital methods. *Public Understanding of Science*, *21*(7), 796–812. https://doi.org/10.1177/0963662510387558

Venturini, T., Bounegru, L., Gray, J., & Rogers, R. (2018). A reality check(list)

for digital methods. *New Media & Society*, 1–23. https://doi.org/10.1177/1461444818769236

Venturini, T., Bounegru, L., Jacomy, M., & Gray, J. (2017). How to Tell Stories with Networks: Exploring the Narrative Affordances of Graphs with the Iliad. In M. T. Schäfer & K. van Es (Eds.), *The Datafied Society: Studying Culture through Data* (p. 13). Amsterdam: Amsterdam University Press.

Venturini, T., Jacomy, M., Bounegru, L., & Gray, J. (2018). Visual Network Exploration for Data Journalists. In S. A. I. Eldridge & B. Franklin (Eds.), *The Routledge Handbook to Developments in Digital Journalism Studies* (p. 19). London: Routledge.

Venturini, T., Jacomy, M., & Pereira, D. C. (2015). *Visual Network Analysis* (Working Paper). Paris: Sciences Po. Retrieved from http://www.medialab.sciences-po.fr/publications/visual- network-analysis

Venturini, T., & Latour, B. (2009). The Social Fabric: Digital Traces and Quali-Quantitative Methods. In *Proceedings of Future En Seine 2009* (pp. 87–101). Paris: Editions Future en Seine.

Venturini, T., & Munk, A. K. (forthcoming). *Controversy Mapping A Field Guide through Actor-Network Theory and Digital Methods*.

Venturini, T., & Rogers, R. (forthcoming). The Cambridge Analytica Affair: the Crisis of API-Based Research and How This May Be a Good Thing. *Digital Journalism*, 7.

Vlist, F. N. van der. (2013). *A Technicity of Collaboration: Coordinating Distributed Software Development on GitHub*. University of Amsterdam, Amsterdam.

Vosoughi, S., Roy, D., & Aral, S. (2018). The Spread of True and False News Online. *Science*, *359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Wardle, C., & Derakhshan, H. (2017). *Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making* (Council of Europe report No. DGI(2017)09). Strasbourg, France: Council of Europe. Retrieved from https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c

Watts, D. J. (2004). The "New" Science of Networks. *Annual Review of Sociology*,

30(1), 243–270. https://doi.org/10.1146/annurev.soc.30.020404.104342

Watts, D. J., & Strogatz, S. H. (1998). Collective Dynamics of 'Small-World' Networks. *Nature*, *393*(6684), 440–442. https://doi.org/10.1038/30918

Weeks, M. (2012). Toward an Understanding of Online Community Participation through Narrative Network Analysis. In H. Li (Ed.), *Virtual Community Participation and Motivation: Cross-Disciplinary Theories* (pp. 90–102). London: IGI Global. Retrieved from https://www.igi-global.com/book/virtual-community-participation-motivation/59749

Weltevrede, E., & Borra, E. (2016). Platform Affordances and Data Practices: The Value of Dispute on Wikipedia. *Big Data & Society*, *3*(1), 2053951716653418. https://doi.org/10.1177/2053951716653418

Weltevrede, E., Helmond, A., & Gerlitz, C. (2014). The Politics of Real-time: A Device Perspective on Social Media Platforms and Search Engines. *Theory, Culture & Society*, *31*(6), 125–150. https://doi.org/10.1177/0263276414537318

Weltevrede, E. J. T. (2016). *Repurposing Digital Methods: The Research Affordances of Platforms and Engines*. University of Amsterdam, Amsterdam. Retrieved from https://dare.uva.nl/search?identifier=aaaa9bb3-8647-41df-954c-2bb1e9f15d77

West, S. M. (2017). Data Capitalism: Redefining the Logics of Surveillance and Privacy. *Business & Society*, 1–22. https://doi.org/10.1177/0007650317718185

Witschge, T., Anderson, C. W., Domingo, D., & Hermida, A. (Eds.). (2016). *The SAGE Handbook of Digital Journalism*. London: SAGE Publications.

Woolley, S. C., & Howard, P. N. (Eds.). (2018). *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media*. Oxford: Oxford University Press.

Zelizer, B. (2004). *Taking Journalism Seriously: News and the Academy*. Thousand Oaks, CA. https://doi.org/10.4135/9781452204499

Zuiderveen Borgesius, F. J. (2014). *Improving Privacy Protection in the Area of Behavioural Targeting*. University of Amsterdam. Retrieved from http://www.ssrn.com/abstract=2654213

Thesis Summary

News Devices: How Digital Objects Participate in News and Research

In this thesis I examine how digital objects participate in news work and research. I focus on the interactions between digital objects, news work and journalism research because news is increasingly taking place in and through a variety of digital devices, from websites, to search engines, online platforms, apps, bots, web analytics, data analysis and visualisation tools. These devices are also increasingly used as resources in digital research. Both these developments are often met with both excitement and concern and their implications are yet to be fully understood.

To address these interactions, I propose the news device research approach. This approach combines STS-inspired socio-material approaches from news and journalism research and device-centred perspectives from digital social and media research. Inspired by Muniesa, Millo and Callon's notion of market devices (2007), news devices refers to objects that participate in news, and the news device approach refers to research outlooks and techniques that treat news work as co-produced through interactions with digital devices, and that capitalise on the medium-specific analytical affordances of these devices to study aspects of both news and digital devices. That is to say, the notion of news devices aims to draw attention not only to how digital devices participate in news work but also to what they may afford for journalism research. In relation to the former, what is important in this project is not the recognition *that* the material participates in news making (this is something that has been recognised for a long time). The aim is rather to empirically explore *when* and *how* it participates and with *what consequences* for the situations in which it is embedded. In relation to the latter, this dissertation does not assume that digital devices straightforwardly and unproblematically hold a promise for news research. But it acknowledges that, given that news and its analysis are increasingly occurring through digital devices, it is important to empirically examine how these devices can be configured to support ways of knowing the news in journalism research.

I illustrate the news device approach through three case studies. Each of these examines different ways in which the digital matters in news and news work

and suggests research techniques by means of which these differences can be attended to. As far as areas of news work are concerned, the case studies cover the participation of digital devices in familiar journalism areas such as that of news stories, as well as less visible ones such as news software development, and the business aspects of news. As far as digital objects are concerned, I focus on three: network diagrams, the code sharing platform GitHub and web trackers. I explore their intersections paying attention to how they shape news work but also to how they shape relations between news and other phenomena and domains, from digital visual culture, to commercial online platforms, the online advertising and marketing industries, and "fake news" or junk viral content production.

In **Chapter 2** I discuss the two bodies of work that I draw on and contribute to, namely STS-inspired news and journalism research and device-centred approaches to digital social and media research. I start developing the notion of the news device as a way to think through how these combined approaches can be brought to bear on the study of digital news making, circulation and use, and the relations and practices that underpin these areas. I focus on these bodies of work because of their widely recognised, rich and nuanced approach to the mutual shaping of technologies and practices, drawing particularly on actor-network-theory (ANT). A number of insights from socio-material approaches in journalism research inform the news device approach, from the sensibility towards the materiality of digital news work, to approaching this work as socio-technical processes, and providing suggestions for aspects of news work in need of further scrutiny. Device-centred approaches from digital social and media research help to further specify and advance this approach, namely by drawing attention to how the digital offers not only sites through which news work can play out, but also to how it can offer medium-specific means to study both news and digital devices. I draw on approaches and insights from digital sociology, software studies, platform studies and research techniques known as digital methods (Rogers, 2013) or interface methods (Marres & Gerlitz, 2015). I end with a discussion of considerations that informed the development of my case studies as well as final aspects of research design.

Chapter 3 focuses on the network diagram as a news device. More specifically, I examine how network graphs, the de facto diagram type of the digital age, make a difference to journalistic narratives about collective life. I examine the narrative affordances of network diagrams, i.e. their capacities to elicit narrative readings in combination with other elements of the journalistic story, through multimodal analysis and approaches from the visual semiotics of network graphs. Although neither representative nor comprehensive, the findings of this chapter suggest that there are patterns in the narrative reading of networks. I describe five recurring narrative readings that the material affordances of network diagrams cue in a collection of journalism pieces. These include: exploring associations around single actors, detecting key players, mapping alliances and oppositions, exploring the evolution of associations over time, and revealing hidden ties. The chapter can be seen as a contribution towards elucidating the meaning-making capacities of these defining knowledge-making devices in the digital age. This has been done by contributing towards an analytical approach to the narrative meaning of networks as well as towards a vocabulary of narratives that networks evoke in journalistic stories. Methodologically, the contribution of this chapter is to illustrate that existing analytical approaches such as multimodal analysis are well suited to support research that takes a news device approach. While this chapter treats the network diagram as an object of research, in Chapter 5 the network diagram becomes part of the visual network exploration method. The double problematisation of the network as object of study and method is thus illustrated across the two chapters. Moreover, the integration of multimodal analysis in the visual network exploration method used in Chapter 5 also illustrates the continuity between research techniques informed by medium-specificity and more established research methods.

Chapter 4 focuses on the online platform as a news device, and more specifically the code sharing platform GitHub. It is a contribution to platform studies, the platformisation of news and to the study of coding and software development in journalism. In the first part of the chapter I analyse GitHub's user interface features, help pages, its development blog and tech press to

understand how GitHub platformises coding, i.e. how it configures coding work into an economically valuable contribution to the platform ecosystem by making participation in the platform traceable and calculable through its technical infrastructure. Drawing on van Dijck (2013) I propose to understand coding on GitHub as connective coding. Unlike the notion of social coding, which draws attention to how GitHub enables connectedness and networking within the software development community, connective coding draws attention to how public coding, developer profiles and behaviours are simultaneously converted into asset streams to be variously capitalised by the platform and its ecosystem. In the second part of the chapter I explore how networked code repositories can be configured into research devices to understand aspects of coding practice in a collection of news code repositories. I examine how date stamps and repository collaboration and popularity metrics can be configured to profile types of platform work, the temporality of coding work and how journalism code is valued on GitHub. I find that journalism coding does not always keep up with platform values (such as its update culture) and that the platform's dominant programming publics shape the valuation of domain-specific journalistic production on the platform. This analysis further nuances the characterisation of platform practices as connective coding by showing that, just as platforms as not fixed or stable arrangements, social practices are also not uniform materials to be platformised but fluctuating, diverse and variable. Methodologically, I make a contribution by exploring how the platform can be configured for news research. I illustrate techniques by means of which researchers can use the platform's analytical capacities at the same time as trying to push back against their logics to align them with their own research questions and interests.

Chapter 5 takes the "fake news" scandal as an occasion to focus on web trackers as news devices and to address the problematic relationship between the online advertising and marketing industries and two forms of digital cultural production: news and viral junk content. To do so it draws on audience economics and socio-material approaches to digital media to study audience commodification from the point of view of web tracking devices. I focus on web trackers as audience marketplace devices because in the context

of the "post-exposure audience marketplace" (Napoli, 2011) in which advertising-supported digital cultural production is increasingly embedded, digital objects such as cookies and other web tracking devices play an increasingly prominent role. To explore how audience commodification can be approached from a digital device perspective, I extract the third-party tracking mechanisms embedded in a collection of pages from mainstream and junk news sites and examine how these devices enact audience marketplace configurations. This analysis provides a window into the audience marketplace practices in which legal initiatives such as the EU General Data Protection Regulation (GDPR) are aiming to intervene. I show that tracking infrastructures are shaped by cultural production dynamics and illustrate a few such audience marketplace configurations. These range from professionalised configurations where tracking is intensified and customised through specialised services and large numbers of intermediaries, to less complex configurations specific to less professionalised publishing operations, which rely on fewer tracking services, most prominently on social media and ad networks. As far as the configurability of web tracking detection for news research is concerned, I document how the technique is shaping the picture of the phenomenon obtained and the difficulties that come with such a research technique due to the instability of the object of study and due to technical aspects of how tracking is detected.

I **conclude** this dissertation by recapitulating the key contributions of this research project and providing some final reflections on their opportunities, challenges and implications for the study of digital news and journalism, as well as some research directions to further explore. I start by revisiting the contributions made through the news device approach and its three empirical applications: the study of networks as storytelling devices, of GitHub as a connective coding device, and of web trackers as audience marketplace device. I discuss what distinguishes this approach from other approaches to the impact of digital technologies on journalism. I also suggest that the promise of news device perspectives is different from that of big data or computational approaches to social and media studies. Unlike these, its strength consists in blending critical research with repurposing the analytical modes inscribed in

devices to understand their interplays with issues and practices. As far as implications for news and journalism research are concerned, I suggested that, while the approach can offer new sites and research techniques to address important questions about digital news and journalism, its implications go further than that. This is because the conceptual and methodological outlooks that underpin it also affect the questions asked and the research problems that are formulated. I consider this interference as positive as it may open up possibilities to address what in Chapter 2 was discussed as perceived blind spots at work in news and journalism research. I also consider it positive as it may be seen as a way to open up meaningful dialogues between the journalism research community and other research communities. I suggest that embracing this approach also involves embracing its particular ways of addressing challenges, which may further unsettle how news and journalism are studied. I end by articulating a present but underdeveloped thread that ran through this dissertation into a direction that future research might consider. It refers to participatory approaches to inquiries into and with digital devices that would see stakeholders affected by such devices, such as journalists, meaningfully participate in the configuration of research inquiries.

Nederlandstalige Samenvatting

News Devices: Hoe Digital Objects Een Rol Spelen in Nieuws en Onderzoek Naar Nieuws

In dit proefschrift bestudeer ik hoe digital objects deel uitmaken van journalistiek en een plaats hebben verworven binnen journalism studies. Centraal staat de wisselwerking tussen digital objects, journalistiek en journalism studies omdat nieuws steeds meer digitaal vorm krijgt en gebruik maakt van digital devices zoals websites, zoekmachines, online platformen, apps, bots, webstatistieken, data-analyse en visualisatietools. Deze instrumenten worden ook steeds vaker onderdeel van digital research. In beide gevallen wordt deze evolutie zowel met enthousiasme als met bezorgdheid onthaald, en moet een getrouw beeld van de eigenlijke implicaties ervan nog gevormd worden.

Om deze wisselwerking te onderzoeken vertrek ik vanuit een news device research approach. Deze koppelt op STS geïnspireerde socio-materiële benaderingen uit journalism studies aan instrumentele perspectieven zoals gehanteerd in digital media studies en digital social research. Het begrip 'news device' is geïnspireerd op het begrip 'market device' van Muniesa, Millo en Callon (2007) en verwijst naar objecten die onderdeel zijn van nieuws. De benadering vanuit news devices verwijst naar perspectieven die journalistiek aanzien als een co-productie voortspruitend uit deze interactie. De mediumspecifieke affordances van deze devices zijn aanzetpunt voor onderzoek naar nieuws en digital devices. News devices hebben dus niet enkel betrekking op de manier waarop ze ingezet kunnen worden binnen journalistiek, maar evenzeer hoe ze een bijdrage kunnen leveren aan onderzoek binnen journalism studies.

In dat eerste perspectief zal het in dit proefschrift niet zozeer gaan over de vaststelling dat materialiteit onderdeel is van de journalistieke praxis; dat is al lange tijd erkend. Doelstelling is eerder om empirisch na te gaan wanneer, hoe, waar en met welke gevolgen deze co-productie zich voltrekt. Wat het tweede perspectief betreft erkent dit proefschrift dat, vermits nieuwsproductie en journalism studies in toenemende mate digitale instrumenten gebruiken, het

van belang is om een duidelijk en empirisch ondersteund beeld te verkrijgen van hoe deze instrumenten kunnen worden geconfigureerd op een dusdanige manier dat ze news research ondersteunen.

De news device approach zal ik illustreren aan de hand van drie casestudies. Deze verduidelijken de verschillende wijzen waarop digitale materie een impact heeft op journalistiek en op journalism studies. De casestudies reiken ook methoden aan om deze aanwezigheid in kaart te brengen en te onderzoeken. De case studies tonen aan hoe digital devices invloed hebben op dagelijkse journalistieke praktijken zoals storytelling. Ze tonen ook aan dat er minder zichtbare impact is, zoals de ontwikkeling van specifieke nieuwssoftware en op de bedrijfsmatige aspecten van mediabedrijven.

In dit proefschrift richt ik me op drie digital objects: netwerkdiagrammen, GitHub als codesharingplatform en webtrackers. Ik verken waar deze digital objects raakpunten hebben, hoe ze news practices vorm geven, hoe ze bijdragen aan domeinen zoals de digitale beeldcultuur, commerciële online platformen en economische aspecten van mediabedrijven. Ook is er aandacht voor de productie van virale junk content en 'fake news'.

In **Hoofdstuk 2** bespreek ik het onderzoeksveld waaruit ik put en waartoe ik een bijdrage wil leveren. Hier bespreek ik de op STS geïnspireerde benadering van journalism studies en de instrumentele bijdrage van digital social studies en digital media studies. Ik ontwikkel het begrip 'news devices' en reflecteer op hoe beide gecombineerd kunnen worden ingezet in onderzoek naar de productie, distributie en consumptie van digitaal nieuws en naar digital news practices. Deze werkwijze biedt een rijk en genuanceerd perspectief op de samenhang tussen technologie en practices door zich te beroepen op de Actor-Netwerktheorie (ANT). De news devices benadering baseert zich op een groot aantal inzichten die zijn voortgekomen uit socio-materiële benaderingen binnen journalism studies. Deze gaan van bewustzijn van materialiteit in digital news practices naar journalistiek als verbonden met socio-technologische processen, alsook specifieke aspecten van journalistieke practices.

De instrumentele perspectieven uit digital social research en uit digital media studies versterken de news devices approach door aandacht te hebben voor de verschillende manieren waarop het digitale perspectief analysemogelijkheden aanreikt voor journalistieke practices maar ook mediumspecifieke parameters naar voren schuift voor onderzoek naar digitaal nieuws. Hierbij maak ik gebruik van benaderingen en inzichten uit de digital sociology, uit software studies, platform studies zoals deze begrepen worden onder 'digital methods' (Rogers, 2013) of 'interface methods' (Marres & Gerlitz, 2015). Aan het einde van dit hoofdstuk bespreek ik de overwegingen van waaruit gekozen werd voor de case studies en de manier waarop ze in dit proefschrift onderzocht werden.

Hoofdstuk 3 behandelt het netwerkdiagram als news device. Ik onderzoek hoe network graphics, het 'de facto-diagramtype' van het digitale tijdperk, vorm geeft aan journalistieke narratie. Ik onderzoek de narratieve affordances van deze diagrammen via multimodale analyse en grafiekensemiotiek. De bevindingen van dit hoofdstuk, hoewel slechts indicatief en niet representatief, wijzen op het bestaan van patronen in narratieve lezingen van netwerken. Ik beschrijf vijf terugkerende lezingen die de materiële affordances van netwerkdiagrammen in journalistieke content oproepen. Het gaat dan om het verkennen van connecties van individuele actoren, het bepalen van de hoofdrolspelers, van bondgenootschappen en allianties, en de manier waarop deze verschuiven doorheen een tijdsframe. Dit hoofdstuk wil meer duidelijkheid verschaffen over de manier waarop deze instrumenten uitgroeien tot instrumenten van betekenisgeving. Zo omvat het hoofdstuk een analytische benadering van netwerken als vertellingen en de manier waarop deze een lezing aanreiken.

In methodologisch perspectief wenst dit hoofdstuk bij te dragen aan het begrip dat een news device benadering prima kan samengaan met andere analytische benaderingen, zoals multimodale analyse. Waar in dit hoofdstuk het netwerkdiagram centraal staat in mijn analyse, zal het in hoofdstuk 5 een onderdeel vormen van de visual network exploration method. Op die manier wordt het netwerkdiagram dubbel aangepakt, enerzijds als onderzoeksobject en

anderzijds als methode. We tonen aan dat er continuïteit is tussen onderzoeksmethoden die geïnspireerd zijn door mediumspecificiteit en andere meer gevestigde empirische methoden door multimodale analyse een plaats te geven binnen de netwerk-verkenningsmethode zoals uitgewerkt in hoofdstuk 5.

Deze analyse brengt verdere nuances aan in het beschrijven van platformpraktijken als connectief programmeren door erop te wijzen dat sociale praktijken, net als platformen, niet stabiel, vast en uniform, maar juist divers zijn, variëren en fluctueren. Ik lever een methodologische bijdrage door te onderzoeken hoe het platform kan worden afgesteld voor onderzoek naar nieuws. Daarbij reik ik technieken aan waarmee onderzoekers de analytische capaciteiten van het platform kunnen benutten, en tegelijkertijd de logica ervan kunnen ombuigen en afstellen naar hun eigen onderzoeksvragen- en interesses.

Ik onderzoek hoe datering van gebruik en vastgelegde gegevens van samenwerking en waardering kunnen worden gebruikt voor onderzoek naar de wezenskenmerken van platform-activiteit. Deze gegevens illustreren de vluchtigheid van programmeerwerk en de waardering voor verschillende onderdelen van journalistieke codes op GitHub. Ik toon aan dat specifieke platformwaarden zoals de updatecultuur een spanningsveld veroorzaken met journalistieke programmeerwaarden. Ook laat ik zien dat het vooral het programmeur-publiek is dat vorm geeft aan de waardering van journalistieke productie op het platform. Deze beschrijving nuanceert het beeld van platformpraktijken als connective coding door te laten zien dat deze praktijken divers, variërend en fluctuerend zijn. Ik zal een methodologisch instrumentarium aanreiken waarmee onderzoekers kunnen recht doen aan de specifieke kenmerken van de journalistieke platformen.

Hoofdstuk 5 vertrekt vanuit de notie van fake news. We analyseren webtrackers als news devices en bestuderen de problematische relatie met de verdienmodellen van digitale journalistieke contentproductie. Ik maak daarbij gebruik van tracking instrumenten om na te gaan hoe audience economics verbonden zijn met de commercialisering van journalistiek en virale junk

praktijken.

Ik analyseer webtrackers als onderdeel van de post-exposure audience marketplace (Napoli, 2011), waar de advertentie-gestuurde digitale contentproductie zich steeds dieper nestelt. Om digital methods te hanteren in het onderzoek naar deze commercialisering extraheer ik de trackingsmechanismen van derde partijen die ingesloten zitten in een aantal pagina's van mainstream- en ook van junknews websites. Ik zal nagaan hoe deze onderliggende instrumenten de publieksmarkt beïnvloeden.

Mijn analyse zal daardoor ook verbonden worden met juridische elementen, zoals de Algemene verordening gegevensbescherming (AVG) waarmee de EU de praktijken op de publieksmarkt tracht te reguleren.

Om aan te tonen hoe sterk de tracking devices verbonden zijn met de culturele betekenisproductie, zal ik in kaart brengen welke vormen van instrumenten gehanteerd worden. Zo zijn er geprofessionaliseerde configuraties waarbij gespecialiseerde diensten middels een groot aantal tussenpersonen tracking intensief en op maat mogelijk maken. Daarnaast zijn er ook minder complexe configuraties toegespitst voor sociale media en advertentienetwerken.

Voor onderzoek binnen journalism studies en voor nieuws is het blootleggen van die tracking practices belangrijk. Ik zal aantonen welke methodologische moeilijkheden overwonnen moeten worden bij gebruik van trackingdetectie, die gelieerd zijn met technische aspecten van onderzoek enerzijds en met de veranderlijkheid en instabiliteit van het onderzoeksobject anderzijds.

Ik sluit dit proefschrift af met de belangrijkste conclusies van de verschillende onderzoeksonderdelen en reflecteer op de mogelijkheden, uitdagingen en implicaties ervan voor toekomstig onderzoek naar digital news en digital journalism.

Ik reflecteer op de bijdrage die de news devices benadering kan bieden binnen journalism studies en de drie case studies als empirische vertaling ervan. Ik

bespreek hoe de news devices approach zich verhoudt tot andere benaderingen over digital journalism en ik vergelijk haar mogelijkheden met die van social media studies, critical research en computational approaches.

De conceptuele en methodologische opvattingen die de news devices approach sturen hebben invloed op de onderzoeksvragen voor toekomstig onderzoek en openen deuren voor het aanpakken van blinde vlekken die nog in de huidige stand van onderzoek aan te duiden zijn.

De news devices benadering levert ook een aantal opportuniteiten als vertrekpunt voor dialoog met andere onderzoeksdisciplines.

Het blijft een uitdaging voor de verdere vormgeving van onderzoek naar nieuws en journalism studies waarbij ook journalisten als belangrijkste stakeholders binnen nieuwsproductie moeten en kunnen betrokken worden.

List of Tables

List of Figures

List of Publications and Presentations Based on This Doctoral Research

Publications

Gray, J., **Bounegru, L.**, & Venturini, T. (forthcoming). 'Fake News' as Infrastructural Uncanny. *New Media & Society*.

Gray, J., & **Bounegru, L.** (2019). What a difference a dataset makes? Data journalism and/as data activism. In J. Evans, S. Ruane, & H. Southall (Eds.), *Data in Society: Challenging Statistics in an Age of Globalisation*. Bristol: The Policy Press. https://doi.org/10.5281/zenodo.1415450

Bounegru, L., Gray, J., Venturini, T., & Mauri, M. (2018). *A Field Guide to 'Fake News' and Other Information Disorders*. Amsterdam: Public Data Lab. Retrieved from https://papers.ssrn.com/abstract=3097666

Gray, J., Gerlitz, C., & **Bounegru, L.** (2018). Data Infrastructure Literacy. *Big Data & Society*, *5*(2), 1–13. https://doi.org/10.1177/2053951718786316

Venturini, T., Jacomy, M., **Bounegru, L.**, & Gray, J. (2018). Visual Network Exploration for Data Journalists. In S. A. I. Eldridge & B. Franklin (Eds.), *The Routledge Handbook to Developments in Digital Journalism Studies*. London: Routledge.

Venturini, T., **Bounegru, L.**, Gray, J., & Rogers, R. (2018). A reality check(list) for digital methods. *New Media & Society*, 1–23. https://doi.org/10.1177/1461444818769236

Bounegru, L., Venturini, T., Gray, J., & Jacomy, M. (2017). Narrating Networks: Exploring the Affordances of Networks as Storytelling Devices in Journalism. *Digital Journalism*, *5*(6), 699–730. https://doi.org/10.1080/21670811.2016.1186497

Venturini, T., **Bounegru, L.**, Jacomy, M., & Gray, J. (2017). How to Tell Stories with Networks: Exploring the Narrative Affordances of Graphs with the Iliad. In M. T. Schäfer & K. van Es (Eds.), *The Datafied Society: Studying Culture through Data*. Amsterdam: Amsterdam University Press.

Gray, J., **Bounegru, L.**, Milan, S., & Ciuccarelli, P. (2016). Ways of Seeing Data: Towards a Critical Literacy for Data Visualizations as Research Objects and Research Devices. In S. Kubitschko & A. Kaun (Eds.), *Innovative Methods in Media and Communication Research* (pp. 290–325). London: Palgrave Macmillan.

Gray, J., Lämmerhirt, D., & Bounegru, L. (2016). Changing What Counts: How Can Citizen-Generated and Civil Society Data be used as an Advocacy Tool to Change Official Data Collection?.

Presentations

The Infrastructural Uncanny and the Social Life of Junk Viral News Online. Presentation at 'Shadow Economies of the Internet' workshop, Stockholm University, Sweden, 20 September 2018.

The Social Life of "Fake News" Online. Guest lecture and workshop at the University of Warwick, U.K., 9 May 2018.

Some Provocations About "Fake News". Invited lecture at 'Inventive Methods for Media Research' lecture series, University of Siegen, Germany, 16 January 2018.

Following Concerning Facts Online: The Case of False Viral News and Other Information Disorders. Keynote at the Digital Methods Winter School 2018, University of Amsterdam, Netherlands, 8 January 2018.

Fake News in Digital Culture. Invited talk at the Department of Digital Humanities, King's College London, 11 October 2017.

Fake News in Digital Culture. Invited talk at the Institute for Policy Research symposium on Politics, Fake News and the Post-Truth Era, University of Bath, 14 September 2017.

A Field Guide to Fake News. Launch event panel at the International Journalism Festival, Perugia, Italy, 7 April 2017 (with Jonathan Gray, Michele Mauri, Angeles Briones, Claire Wardle and Craig Silverman).

Fake News, Algorithmic Accountability and the Role of Data Journalism in the Post-Truth Era. Invited talk at the workshop 'How Can Public Interest Journalism Hold Algorithms to Account?', University of Cambridge, 23 March

2017 (with Jonathan Gray).

What Can Data Journalists and Digital Humanists Learn from Each Other? Invited talk at the Digital Humanities + Data Journalism Symposium, University of Miami, US, 30 September 2016 (with Jonathan Gray).

Data Journalism and the Remaking of Data Infrastructures. Invited talk at the Evidence and the Politics of Policymaking: Where Next? Symposium, Institute for Policy Research and Centre for Development Studies, University of Bath, UK, 14 September 2016.

Data Infrastructure Literacy: Reshaping Practices of Measurement, Monitoring and Evidence. Conference paper at 4S/EASST Conference Barcelona 2016, Spain, 2 September 2016 (with Jonathan Gray and Carolin Gerlitz).

Doing Digital Methods. Some Recent Highlights from Winter and Summer Schools. Invited talk at the Digital Methods Summer School, Digital Methods Initiative, University of Amsterdam, Netherlands, 27 June 2016.

Ways of Studying, Using and Intervening in Data Infrastructures. Invited seminar at Sciences Po Paris médialab, France, 19 April 2016 (with Jonathan Gray and Carolin Gerlitz).

The Rise of Data Journalism: The Making of Journalistic Knowledge through Quantification. Invited lecture at Sciences Po Paris, France, 5 April 2016 (with Jonathan Gray).

Ways of Seeing Data: Towards a Critical Literacy for Data Visualisations as Research Objects and Research Devices. Paper at the Digital Methods Winter School Conference, University of Amsterdam, Netherlands, 14 January 2016 (with Jonathan Gray).

Towards a Literacy for Data Infrastructures. Paper at the Digital Methods Winter School Conference, University of Amsterdam, Netherlands, 14 January 2016 (with Jonathan Gray and Carolin Gerlitz).

What Does a Good Digital Methods Project Look Like? Some Highlights from DMI Winter and Summer School 2015. Invited talk at the Digital Methods Winter School 2016, University of Amsterdam, Netherlands, 11 January 2016.

From Telling Stories with Data to Telling Stories with Data Infrastructures: Repurposing Digital Methods and the Data Sprint for Data Journalism. Invited lecture at the University of Amsterdam, Netherlands, 7 January 2016.

Doing Social and Political Research in a Digital Age: An Introduction to Digital Methods. Invited lecture at the University of Ghent, Belgium, 4 December 2015.

Doing Social and Political Research in a Digital Age: An Introduction to Digital Methods. Invited lecture and seminar at the National Centre of Competence in Research: Challenges to Democracy in the 21st Century, University of Zurich, Switzerland, 5-6 November 2015.

Redistributing Journalism: Journalism as a Data Public and the Politics of Quantification in the Newsroom. Paper given at the Data Power Conference, University of Sheffield, UK, 22 June 2015.

Storytelling with Data Visualisation: An Introduction. Invited workshop given at the Amsterdam School of Communication Research (ASCoR), University of Amsterdam, Netherlands, 10 June 2015 (with Jonathan Gray).

Using Data for Science Journalism. Invited talk and workshop at the International School of Science Journalism, Erice, Italy, 10 May 2015 (with Jonathan Gray).

Follow the Networks: Open Data and Digital Methods for Journalism. Invited lecture at University of California Berkeley, US, 20 April 2015 (with Jonathan Gray).

Improving the Coverage of Complex Issues with Data Journalism and Digital Methods. Invited talk at the BBC Data Day, London, UK, 12 November 2014.

An Epistemological Experiment: Issue Mapping, Data Journalism and the Public Understanding of Complex Issues. Invited lecture at the Department of Media and Culture Studies, Utrecht University, Netherlands, 27 October 2014 (with Jonathan Gray).

What Actor-Network Theory (ANT) and Digital Methods Can Do for Data Journalism Research and Practice. Invited lecture at Ghent University, Belgium, 21 October 2014.

Mapping Issues with the Web: An Introduction to Digital Methods. Invited talk at Columbia University, New York, US, 23 September 2014 (with Jonathan Gray).