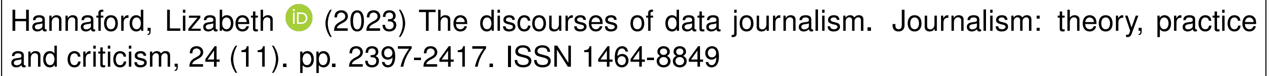
Manchester Metropolitan University
e-space
Manchester Metropolitan University's Research Repository

Please cite the Published Version


Hannaford, Lizabeth (2023) The discourses of data journalism. Journalism: theory, practice and criticism, 24 (11). pp. 2397-2417. ISSN 1464-8849

DOI: https://doi.org/10.1177/14648849221107803

Publisher: SAGE Publications

Version: Published Version

Downloaded from: https://e-space.mmu.ac.uk/629840/

Usage rights: Creative Commons: Attribution 4.0

Additional Information: This is an open access article which appeared in Journalism, published by SAGE Publications.

Enquiries:
If you have questions about this document, contact openresearch@mmu.ac.uk. Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines)

Check for updates

Original Article

The discourses of data journalism

Journalism
2022, Vol. 0(0) 1–21
© The Author(s) 2022

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/14648849221107803
journals.sagepub.com/home/jou

Mrs Liz Hannaford, MSc
Manchester Metropolitan University, Manchester, UK

Abstract

Becoming a data literate, technologically competent journalist is represented as a desirable goal that will benefit the individual, the industry and society as a whole. Data journalism skills are increasingly being taught in journalism programmes around the world. This article applies Foucault's distinctive conceptualisation of discourse to critically examine data journalism as constructed in the 'talk' of its most visible pioneers. The analysis is driven by three distinctive aspects of Foucault's theory of discourse – power, knowledge and materiality. Using these tools, I investigate how data journalism knowledge is produced, the practices that reinforce it and the strategic power relations it conceals. I argue that data journalism draws on four discourses – journalism, technology, enterprise and citizenship – and wraps itself in the power relations embedded in these prestigious discourses. I argue that there is a political imperative for journalism educators to examine these power relations because material injustices along race, gender, class lines are built into them and have consequences for our students and society.

Keywords

Computational/data journalism, education, discourse analysis, critical theory, foucault

Introduction

This paper analyses the discursive practices that have constructed data journalism as an accepted way to talk about quantitative, computational approaches to newswork (Coddington, 2019). I analyse the expert talk of the field's pioneers using Foucault's (1981) distinctive concept of discourse to uncover the power relations woven into data journalism's optimistic language. I argue that there is a political imperative for journalism

Corresponding author:
Liz Hannaford, Languages, Information and Communications, Manchester Metropolitan University, Grosvenor East Building, Cavendish Street, Manchester M15 6BG, UK.
Email: L.Hannaford@mmu.ac.uk

educators to examine these power relations because material injustices along race, gender, class lines are built into them and have consequences for our students and society. Drawing on Foucault enables us to look beyond data journalism as simply a neutral set of skills that exists separate from the world. It was not self-evident that a desk in the newsroom would be dedicated to the analysis of numbers or that coders might sit at this desk rather than in the IT department. It was not inevitable that this would be seen as exciting, significant journalism worthy of its own awards (Ojo and Heravi, 2018). Nor was it inevitable that universities would dedicate courses to it (Heravi, 2019). Instead, this paper understands data journalism as a rationalised and contingent category of knowledge shaped by discourse. The effect of discursive practices is to make it virtually impossible for a right-thinking person to think outside of them because the power relations are reproduced so extensively and subtly (Foucault, 1981; Hook, 2001).

This critical approach is well established in other fields of education (Peters et al., 2009) where 'learning' is situated as the embodiment of historically constructed values that determine how we should conduct ourselves (Popkewitz and Brennan, 1998). For example, how is nursing knowledge legitimated (Springer and Clinton, 2015) and how do science textbooks construct particular student subjectivities (Bazzul, 2015).

However, the literature around journalism education has been criticised for its tendency to be theoretically limited, normative and descriptive (Deuze, 2006). Its generalised acceptance of the stability and homogeneity of the industry (Deuze and Witschge, 2018) has led to a narrow focus on socialising students to the routinised practices of the newsroom model (Mensing, 2010) rather than a more critical engagement with how normative ideas of journalism are channelled into university education and whose voices are amplified or marginalised in this process. The scholarship of data journalism education is inevitably even more limited but there are useful reports on the extent of data journalism education globally (Berret and Phillips, 2016 in the USA; Davies and Cullen, 2016 in Australia; Heravi, 2019 globally; Splendore et al., 2016 in 6 European countries; Yang and Du, 2016 in Hong Kong) and the challenges it poses to educators (Berret and Phillips, 2016; Green, 2018; Heravi, 2019). Beyond this, the literature is mainly confined to descriptions of model curricula and 'innovative' pedagogy (Bradshaw, 2018; Green, 2018; Hewett, 2015; Treadwell et al., 2016).

Following an explanation of how I put Foucault's distinctive conceptualisation of discourse to work in this paper, I identify four discourses that are employed by pioneers of data journalism – journalism itself, technology, enterprise and citizenship. I show the power effects of these discourses, the practices that reproduce them and the injustices they obscure. I then offer some thoughts on how this might be useful to journalism educators.

Discourse approaches to data journalism

The narrow approach of the journalism education literature contrasts with developments in journalism studies more broadly where there is growing emphasis on the fragmentation and heterogeneity of newswork in the digital era (Witschge et al., 2019). Kreiss and Brennen (2016) summarise the key themes of recent scholarship as participation, de-institutionalisation, innovation and entrepreneurialism but argue these normative claims

have often been embraced uncritically without attention to who is advantaged and disadvantaged. As data journalism emerged as a distinct field within digital journalism, scholarly interest understandably concentrated on the practical implications - who does what (for example Appelgren and Nygren, 2014; Karlsen and Stavelin, 2014; Royal, 2010) and what counts as 'good' data journalism (Loosen et al., 2017; Ojo and Heravi, 2018). However, a more theoretically-informed direction has evolved (see Coddington, 2019 for an overview) and I focus here on the discourse-oriented approaches.

These approaches focus attention on conventionalised ways of talking about journalism (Borger et al., 2013) and how these discourses constitute material effects. Carlson (2015) uses a framework of metajournalistic discourse to examine journalism as a contextually-embedded cultural practice which must negotiate "moments of contestation" (Carlson, 2015: 352) when taken-for-granted assumptions are challenged and the boundary between acceptable and deviant practices needs to be rearticulated. Journalism's quantitative turn can be understood as just such a "moment of contestation" requiring a discursive response in order to establish authority (Lewis and Waters, 2017).

Research in this tradition locates data journalism in the discourses it produces rather than in its artefacts (De Maeyer et al., 2015). Data journalism is constructed as a reality by social actors and made to seem reasonable and 'true.' Journalists, through a "process of articulation" (Bucher, 2017: 919), strategically join together discourses to legitimise certain ways of thinking about the world. These discourses necessitate the material structure of organisations such as Hacks/Hackers which reproduce the discourses in a mutually-shaping relationship (De Maeyer et al., 2015). Anderson (2018) uses Foucault's concept of genealogy to unpick the historical regimes of meaning-making that enable certain technologies and objects to be adopted by journalists. Borges-Rey (2016) draws on Foucault to show how data journalists justify their accounts of the world by representing numeric infallibility and computational neutrality as unquestionable 'truths.' These discourses constitute a new category of "technologically savvy journalist" (Creech and Mendelson, 2015: 153) who is constructed as an 'ideal' subject.

This paper builds on this literature by operationalising the distinctive aspects of Foucault's conceptualisation of discourse to analyse the expert-talk of pioneering practitioners. This is useful for journalism educators who are shaping professional practices and identities. It helps us to problematise data journalism beyond a skillset to be squeezed into an already tightly-packed curriculum. It makes visible, for example, taken-for-granted ideas about technology and society and how we might wish to challenge these with our students in the classroom.

Foucauldian approach to discourse analysis

For Foucault, discourses "systematically form the objects of which they speak" (Foucault, 2002: 54), rationalising and normalising what it is possible to say and do. But discourse is not purely linguistic or deterministic. It has a historical dimension in that the 'truth' produced by discourse is contingent, underpinned by the socio-political context (Hook, 2001). Discourse for Foucault is always materially embedded in that it produces practices that extend the reach of the discourse and intensify its power effects. At the same time, it

conceals its power relations, putting the discourse seemingly beyond question. Thus, Foucault's conceptualistaion of discourse has three distinctive elements that drive the analysis in this paper - knowledge, power and materiality (Foucault, 1981, 2002; Hook, 2001).

Discourses produce frameworks of legitimate, value-laden knowledge that function as 'truth' and divide what is reasonable to say from what is 'madness' (Hook, 2001: 523). But this knowledge does not exist in a void; it is instrumentalised, producing real effects and institutions that make it 'more true.' For Foucault, discourse is always political - a strategic game in which we desire to be seen as speakers of the 'truth' in competition with other discourses. But this power play must also be obscured so that the truth it underpins is accepted as natural, inevitable and freed from power. For Foucault,

> power is tolerable only on condition that it masks a substantial part of itself. Its success is proportional to an ability to hide its own mechanisms. (Foucault, 1990: 86)

Therefore, successful discourses must strategically ground themselves in that which is already validated and considered natural and common-sense. For example, knowledge of data journalism does not begin out of nowhere. It joins the voices already speaking (Foucault, 1981), the existing systems of knowledge with their established institutional supports, structures and power relations (e.g. the institution of journalism, technology, market economics, liberal democracy). This enables data journalism to compete with other notions of 'good journalism' – and conceal the power relations it appeals to - by cloaking itself in pre-existing truths, values and rationality (Hook, 2001). A Foucauldian analysis will seek to expose these power relations to show how data journalism's pioneers embed themselves in broader power networks and, at the same time, confer power to these existing networks by reproducing their discourses in a positive way. This relies on Foucault's radical conceptualisation of power as a network of relations operating within a discourse or field. So rather than viewing power as repressive and hierarchical (Clegg, 1989), it focuses on how multiple actors over time intensify the power effects of a discourse through their talk and practices (e.g. individuals who run data journalism meet-ups, address conferences, judge data journalism awards).

This external materiality and strategic power are what move Foucault's concept of discourse beyond the purely linguistic (Foucault, 1981). An analysis drawing on Foucault looks at how the discourse is instrumentalised – what does it make people do, what physical spaces are formed and what artefacts does it give meaning to – within a specific historical period (Hook, 2001: 9). Discourses incorporate practices. Practices bring data journalism into being, intensifying its power effects (Hardy and Thomas, 2014). So these extra-textual forms of power relations in practice (e.g. the arrangement of desks in the newsroom, who speaks, what tools are valued) are key to a Foucauldian approach to discourse analysis that emphasises the 'doing' of data journalism. Furthermore, discourse itself has a material context. In this case, data journalism seeks to assert itself as a valid and, indeed, superior form of journalism that solves assumed problems that exist beyond the text. Finally, discourses are 'authored' (Foucault, 1981) by subjects who identify with

the object of discourse, engage in practices that reinforce it and make possible certain subject positions i.e. the data journalist.

The research question derived from this use of Foucault is:

> What power relations do visible pioneers appeal to when they make 'truth' statements about data journalism and in what material practices are these embedded?

Methodology

I now explain how I have applied Foucault's concept of discourse to the expert talk of data journalists. As it emerged as a distinct field with few text books, data journalism's pioneers (Hepp and Loosen, 2019) regularly shared their expertise through public appearances at training courses, conferences, networks (Bounegru and Gray, 2021). In selecting pioneers for this analysis, key starting points were contributors to the first edition of the Data Journalism Handbook (Gray et al., 2012) and the instructors for the European Journalism Centre's 2014 MOOC for Doing Journalism with Data, accessed by over 21,000 people from more than 170 countries (Howard, 2014). The International Journalism Festival is the largest annual media event in Europe and regularly discusses data journalism providing another source of visible experts as did networks and organisations such as The Global Editors' Network, The Tow Center for Digital Journalism and the Nieman Foundation which "shape the field of journalism" (Creech and Mendelson, 2015: 184).

I narrowed my text selection to interviews, talks and discussions in which the communicative purpose (Fairclough, 2003) is to disseminate expert knowledge less formally than in written texts (for analysis of *written* texts see Creech and Nadler, 2018; Powers, 2012). Since visibility was a key requirement, only texts freely available online are included. The timeframe chosen is 2008–2018 which marks the emergence of data journalism as a distinct field and broadly coincides with the first decade of The Guardian's influential Datablog, founded by Simon Rogers in 2009. However, Rogers credits a talk given by Adrian Holovaty at *The Guardian* in 2008 as the inspiration for the Datablog (Gambini, 2019), so it seems relevant to this study to use that talk as the start point for the timeframe, whilst recognising that the discourses should not be considered to have been static in this period but subject to mutations.

These criteria narrowed the selection to 26 texts – listed in the Appendix - each of which was assigned a code. Transcribing the recorded material produced a dataset of over 100,000 words about data journalism which is sufficient to allow for variation and fresh insight whilst remaining manageable for the depth of analysis.

The list of pioneers is inevitably subjective and, in choosing the UK context for study, it is recognised data journalism may be constituted by different discourses in different social and political systems. However, the UK was an early stronghold of data journalism (Stalph and Borges-Rey, 2018) and so offers a historically relevant context. Furthermore, these pioneers' visibility is in part due to their controversial views and therefore the discourses they produce do not represent the entirety of the discourses about data

journalism. However, the regularity with which they are invited to present their views means they become privileged voices within the network.

Discourse analysis is not a single analytical method but rather a field of research encompassing a variety of theoretical, ideological and ontological positions (Wetherell et al., 2001a). Discourse analysis in the Foucauldian tradition highlights the materiality of discourse, excavating and making visible the power relations concealed in truth statements. Repeated close reading of the dataset identified regular themes, arguments and framing which, through an iterative process, were then categorised into four distinct strands that appeal to broader, established discourses.

The paper ends with reflections on the implication of this analysis for journalism educators.

The interlocking discourses of data journalism

I identify four discourses regularly employed by visible pioneers to construct data journalism knowledge. Each of these prestigious discourses has its own rationalised 'truth,' power relations and practices that are used by the speakers to construct and validate data journalism.

1. Journalism
2. Technology
3. Enterprise
4. Citizenship

It is important to note the very term data journalism is explicitly problematised by speakers in these texts as unsatisfactory - "has come to mean everything and nothing" (B2).

Journalism discourse - "It's just journalism"

This discourse constructs journalism as a closed institution with its own stabilised routines and practices (Mayr, 2008) enabling the speakers to claim professional legitimacy and status as members of a valued social group. This is the most conservative of the four discourses in that speakers legitimise data journalism by appealing to the established norms and power relations of traditional Western journalism. Indeed, these speakers predict the modifier "data" should eventually become unnecessary:

> the sexy term that everybody throws around is… - I really do believe - will disappear and it will just be journalism. (J1)

"It's just journalism" becomes a mantra in this discourse as a strategic power move that attempts to shift data journalism from the margins of newsroom practice and assimilate it with the dominant paradigm. Drawing on Foucault's notion of discourse as materially embedded, the physical space occupied by data journalism is shifted from the outsiders'

"corner" (B2), to the centre of the newsroom alongside editorial desks that bestow prestige. Beyond the text, the material context of this discourse is the perceived threat to institutional journalism from the public's lack of trust in news. Data journalism is constructed as the solution by claiming it is a form of news that unfailingly delivers unbiased accuracy and 'truth.'

For speakers of this discourse storytelling is sacrosanct and information must be narrativised in this way in order to have value. It assumes a shared conception amongst journalists of what a 'story' is (Zelizer, 2005) and confers power to journalism as a closed group by reproducing this institutionalised knowledge.

> "People think of data journalism as about producing charts. It's not. That's data visualisation. Data journalism is the process of telling a story with data and getting stories out of those numbers." (A3)

> "You need to have a clear story, plots that evolve, that make for a good reading. You cannot just display data and say, 'Oh, I'm done.'" (I4)

This discourse privileges words and the effort it takes to convert data into narrative. It is thus an appeal to the already rationalised structures of traditional journalism whereby a journalist is an individual who exercises power through generating persuasive interpretations and factual narratives.

As a further power play, this discourse claims it is the natural heir of prestigious investigative journalism (De Burgh, 2000). Investigative journalism is associated with elite news organisations. It is inherently values-driven, based on institutionalised ideas about what constitutes bad behaviour and whose bad behaviour the collective "we" should be concerned about – sport governing bodies and banks are examples found in this dataset. The role of technology in these data investigations, whilst highlighted, is blackboxed (Latour, 1994) as a tool for enacting normalised professional practices:

> "helping the traditional journalists move faster and dig deeper." (J2)

Journalism is assumed to be a serious and impactful pursuit that uses its resources to expose injustices. This claim is instrumentalised in this discourse through time-intensive practices of digital analysis that are evaluated by speakers as "exhaustive," "really hard," "crucial." Thus, the 'truth' of what journalism is - its function and daily rituals – is narrowed as this discourse distances itself from everyday reporting tasks. It wraps itself in the highly prestigious institution of the well-resourced newsroom holding authority to account.

This pioneer discourse places data journalism within a neatly constructed, linear history of journalism that assumes it has evolved inevitably towards its 'true' end point of data journalism. Here, longevity is used to bestow legitimacy and prestige so speakers reach back into history to find exemplars that 'prove' data journalism was always meant to be. Two speakers draw on an example that has become part of the canon of data journalism literature – John Snow's cholera map of 1854. This map normally belongs in the context

of epidemiology but is recontextualised as a journalistic endeavour. We are told it is "an amazing data visualisation" that "literally did change the world" (A3) thus appealing to the established medical discourse to associate data journalism with curing cholera! Another speaker constructs a coherent history with an origin point in the print era, citing Pulitzer-prize winning examples. This enables the speaker to make a power claim that data journalism is not a gimmick invented by data scientists but is integrated into the existing power structures of institutionalised journalism.

The journalism discourse produces knowledge of data journalism as a rigorous form of technology-enabled, factual storytelling. The pioneer speakers construct is as 'better' journalism by strategically positioning it as close as possible to established power networks associated with heroic, value-laden journalism done for the public good. Through this legitimation, it exercises its power to make authoritative knowledge claims that serve the needs of a passive audience. This audience is constructed as a collective with shared interests and concerns about elites. But it is an audience that requires numerical data to be packaged as a narrative that conforms to the established format of 'journalism.' The discourse is materially embedded in that it introduces new terminology and data skills into journalism competency, it occupies space and allocates resources in the newsroom and publishes high-profile articles that intensify the discourse. It conceals its own power claims with appeals to the power network of mainstream journalism which is associated positively here with truth, accuracy, justice. In so doing, it obscures the inequalities built into this power network such as its imposition of Western-centric notions of journalism and its institutionalised ideas of whose injustices merit attention and whose do not. By identifying with this discourse, subjects of data journalism understand themselves as part of a hierarchical profession with normative routines that can be improved with technology. It hints that a journalist whose skillset does not include data literacy could become a 'risk.'

Technology discourse – "We're doing things that a human could never do"

In this discourse, technology is represented as disruptive, resisting the norms of the traditional journalism discourse. It assumes the neutrality of the machine and its ability to free the human from drudgery. The power relations of Big Tech speak through this discourse which is materially embedded in mass computation and datification. From the teleolgical perspective, traditional journalism is assumed to be broken and the technology discourse guarantees to fix it, promising a utopian era in which the 'closed' world of journalism is opened up to new categories of individuals.

Datification - the rendering of our experiences and actions as streams of machine-readable data (Van Dijck, 2014) - is presented here as a sign of progress. Even as elsewhere in society there were calls for protection from the threat of privacy invasion and surveillance (Pilkington, 2015), these threats are backgrounded in these pioneer discourses. Instead, we are encouraged to rejoice in a world in which we are "drowning in data":

> "With organisations releasing records, with them digitising, and producing data every single day, on our phones, on our computers, we tap in and out of the Tube, we walk down the road and CCTV captures us. Every minute of our lives data is being captured, and that's a lot of data, a lot of information. And there are stories after stories after stories buried in that. So the benefits of being able to mine it and to analyse it is an added benefit because that is where stories are, and that's where we need to get to." (J3)

The processes that led to the existence of that data are blackboxed here as part of the natural processes of everyday life. The power relations involved in how "walking down the road" becomes data and why it should be data and whose interests are served by it are obscured in this pioneer discourse by appealing to "benefits" such as progress and truth. Speakers repeatedly refer metaphorically to data as a natural, "real" phenomenon – a raw mineral that can be "mined" or "dug into" or a body of water – a "deluge" - in which we might "drown" but into which we should "dive." The stories are "buried" or "hidden" in the data, waiting to be "found" rather than constructed by the journalist. And because data is presented as pure and natural, the stories 'found' within it are presented as neutral, truthful copies of the world.

This discourse produces new knowledge of what journalism is. It devalues the story format and assumes users "really crave" unmediated access to "raw data:"

> "They want the interpretation and the analysis from people, but they also want the veracity of seeing the real thing, without having it aggregated or put together. They just want to see the raw data." (A1)

"Interpretation" and "analysis" as performed by "people" is contrasted with "veracity" and "seeing the real thing." Users, we are told here, should be engaged not passive and this is reinforced through the practice of interactive tools created by tech-savvy journalists. The design elements and coded algorithms that produce the interactives are represented as neutral rather than constraining.

> [This interactive election map] "combines demographic and past election data with the ability for users to make a choice and deeply engage with the interactive. It's not just telling the user a story, but informing the user by allowing him or her to be part of the story. That, I believe, is when data journalism becomes its most compelling and informative." (F1)

This interactivity is presented as a resistance to traditional journalism's paternalistic gatekeeping. Instead of newsworthiness being defined by powerful institutions, it is reconceptualised as highly localised information that can serve the user's personalised needs:

> "if your local restaurant is inspected and actually passes the inspection, that's completely not newsworthy in the major newspaper sense but it's still newsworthy if you live near it. […] So there's a huge opportunity there for doing, for doing a lot of good." (D1)

The machine is represented as liberating journalists from the drudgery of repetitive journalistic tasks such as reading through documents. Automation, portrayed as a threat to human happiness in other discourses, is seen here as a "beautiful thing" (J2) meaning journalists "don't have to worry about manually doing any sort of labour" (D1). It is assumed to be self-evident that "labour" causes us "worry." A sort of cyborg is envisioned whereby the human essence and technology are entirely compatible (Fisher, 2010) – "we're doing things a human could never do" (J2).

This discourse does not simply describe technology; it entangles journalism with the new, powerful apparatus of technology giants. It confers power to these organisations that collect and control information about us by associating them with benefits, productivity, ease.

> "We run almost everything off of Google Spreadsheets […] because keeping it simple is, you know, arguably more important." (G1)

> "So we made this map just using Google Fusion Tables. It took about half an hour to make." (A3)

The material effect of this discourse is the deep penetration of Google, Microsoft etc. into news work such that they are seen as benevolent co-producers pursuing the same aims as journalists. At the same time, this discourse obscures the power relations that keep journalists dependent on these companies' free, easy-to-use tools by removing the need for advanced skills such as coding. Although this is presented here as liberating and democratising, it potentially disempowers us. For example, Google Fusion Tables was widely used by data journalists to merge and map datasets but was switched off by Google in 2019 and the content that used it disappeared (Heravi et al., 2021).

By masking Big Tech's impact on privacy, the economy and security, this discourse can strategically present data journalism as fixing a broken journalistic culture. Here, there is no neat, linear history. Instead, there is a radical schism separating the traditional journalism of the past from what we are told is the inevitable future. The audience is reconceptualised as civically engaged but fragmented individual users with personalised interests and the means to engage with 'raw data.' A data elite is constructed which excludes the data poor who do not have access to these new digital tools. It produces the subject position of a tech-savvy journalist who resists the rules of traditional journalism.

Enterprise discourse – "We need to learn from Silicon Valley"

This discourse produces knowledge of journalism as a commodity in a competitive market as opposed to a public service necessary for democracy. Work intensification is understood to be an inevitable, positive consequence of technological advancements presenting opportunities for the 'right' individuals to move ahead. "Legacy" news organisations are associated with institutional inertia that cannot adapt to an era of risk and uncertainty. The material context of this discourse is that data journalism can solve the existential crisis of news organisations:

> "There's much more of an argument for what we do because the business model is that this is news worth paying for, that you're trying to give your audience and your readers something exclusive, something they can't get anywhere else, something that is worth subscriptions." (J1)

Because it is assumed that the market can resolve the structural problems facing the news industry, alternative support systems for public service journalism need not be considered. The audience is now made up of individual "subscribers" – not a citizenry that needs to be informed - making choices about which news product is "worth" paying for. These discerning subscribers are assumed to value exclusivity and desire to pay for a prestige product. But the unjust power relations of the market are obscured - the members of society who do not have the means to subscribe are excluded and the divide this creates is not addressed.

When the pioneers use this discourse, they refer to broader discourses that invoke market economics. We are urged to learn from Silicon Valley rather than Fleet Street to find the solutions to the challenges facing news organisations. The many failures littering Silicon Valley are made invisible and instead we are told to admire the efficient strategies of these tech companies that thrive in the changing world. One speaker claims journalism is analogous with video games because both tell complex stories – ignoring the epistemological and social differences between these industries.

> What's the big difference between World of Warcraft and stories in newspapers? In my opinion, it's that World of Warcraft is making tons of money. It's sexy. It's interesting. It's successful. And I'm saying that not because I'm saying that journalism has to become like World of Warcraft. But I'm saying that because if you think about video games, the video industry in the 80s, it was Prince of Persia. Today, you have massive, multiplayer games online. You have games on FaceBook, on Twitter. You have mixed reality games. You have a whole lot of new innovative things. […] I think the video game industry has moved much faster than journalism. (I1)

Organisations are represented as having to be in a constant state of innovation in order to be secure in a precarious, competitive market. "Innovation" blurs into a particular entrepreneurial logic (Creech and Nadler, 2018; Kreiss and Brennen, 2016) defined by "sexy" tech companies and based on the taken-for-granted notion that competition drives innovation.

The speakers don't just talk about enterprise; they engage in practices that reinforce this discourse's 'truth' such as efficiency, monitoring 'traffic', measuring a story's 'performance.' These blur the physical divide between the editorial and business side of news organisations and produce new journalistic roles such as product manager (Royal et al., 2020). It constructs a new journalistic subject who is performance-conscious and should use the language of the market and accounting which, in turn, extends the reach and power effects of this discourse. It commodifies the audience into 'subscribers' which requires a suite of mechanisms to target and count these. The enterprise discourse positions data journalism as the format best suited to capitalise on the market conditions of

journalism by appealing to the privileged power networks of commercial enterprise including Silicon Valley. By associating this with positive values such as success, exclusivity, innovation it masks the material injustices of this system - the white male-dominated culture, the widening technology gap between rich and poor countries and the environmental impact.

Citizenship discourse – "Everything we do is about transparency"

In this discourse, speakers make statements about how journalism should function in a datafied society. The closed newsroom practices of traditional journalism are made problematic and it is taken-for-granted that collaboration with like-minded journalists and audiences across the world will lead to better journalism. This discourse challenges the status quo by valuing openness and collaboration more highly than ownership and control. It enlists existing 'knowledge' about 'transparency' as a self-evident good to increase the standing of journalism in society (Allen, 2008). It is made natural here that data journalism should be driving society towards greater openness.

> "What's the future hold for data journalism? I think it'll only become more transparent, more available, more accessible for more and more people to do. More and more news organisations are realising every day this is what they should be getting involved in and I think… but it's a promoting a greater openness and literacy in society as a whole and that's got to be a good thing." (A4)

By equating "greater openness" with "literacy in society," the speaker appeals to the optimistic neoliberal idea that open data is the same as open culture and will lead to economic growth, better decision-making and innovation. This notion is operationalised in the legal and administrative apparatus of Freedom of Information requests which have become a key part of the data journalist's practice. By linking in with the power network of rights and freedoms, this pioneer discourse seeks to rationalise a techno-heroic role in resistance to the 'secrecy' of governments.

> "If truly every government body and every organisation that is public opened their data, you wouldn't need to do that [Freedom of Information requests] to begin with. And my focus has always been on that, the source of that. I think the fact that FOI is under threat is a travesty and it's absolutely unacceptable that they're threatening it because this is an affront to a public service. This is a right being taken away from citizens." (J1)

The pseudo-transparency of governments and organisations that release data in the closed PDF file format is mocked but, we are told, data journalism has the technical skills to subvert this perceived hypocrisy by converting PDFs into useful, machine-readable formats. This act of subversion is "one of the more noble things we as data journalists can do" (E1) enabling this discourse to derive authority from its ability to produce 'transparency.' This draws on the power relations around the dominant Western paradigm of liberalism, couched in the positive language of individual rights, freedom and access. But

by masking these power relations in the optimistic language of empowerment, it obscures the ways in which open data – and liberalism itself – perpetuates the digital divide. Access to open data is not the same as having the resources to make effective use of it (Gurstein, 2011) and so it potentially empowers those who are already empowered through their education, financial security, able-bodiedness. The hierarchical power and governmental apparatus required to administer 'rights' and the many exclusions, exploitations and contradictions liberalism creates are glossed over.

The traditional journalism discourse is problematised here because of its institutional control which is contrasted with the transparent and participatory orientation of the open-source, hacker community (Baack, 2018) which speakers of this discourse regularly appeal to. Journalism is a problem to be solved and 'hacker,' we are told, is a "term of endearment," (E1) and is "pretty cool" (G1) because of its playful, collaborative approach to civic problem-solving.

The materiality of this discourse is the growth of the meet-up scene where journalists and technologists gather in informal social venues, accompanied by food and drink, to share knowledge and 'rethink' journalism. This intensifies the discourse's claims to be part of a wider movement of technology-enabled civic activism. Speakers engage in practices such as harnessing the collective knowledge and labour of the audience. It is accepted expertise does not rest exclusively in the closed space of the newsroom but is dispersed in the wider world. The small, Paris-based team, OWNI, did not have the resources to compete with The Guardian or the New York Times to find stories in the Wikileaks war logs:-

> "So we decided to be collaborative and to invite users in. So, erm, what we did first was to ask for help. As soon as we decided to go ahead with this story, we sent out messages on Twitter asking for web users to help us. The first thing was to understand the documents and to translate all the NATO jargon into French. So we had about 15 people who volunteer to help us on this Monday morning." (I1)

The citizenship discourse produces knowledge of journalism as an integral part of participatory democratic society, subverting the assumed secrecy of authority. It employs strategies that appeal to the rationalities of Western liberalism and obscures the ways in which rights themselves function as forms of power, constraining the subject and entrenching race, gender and class inequalities. The discourse is materially embedded in the practices of FoI requests and a vibrant meet-up scene that organises spaces and time for journalists and technologists to collaborate. The meet-up scene's physical and online presence intensifies the reach of this discourse by reproducing the notion of a 'community' of public-spirited individuals for whom the boundaries between work and leisure are blurred. The audience is constructed here as an active co-producer of knowledge with the means to make effective use of data. The journalism subject who identifies with this discourse understands themself as a transparency-advocate and civic-minded rights-holder.

Implications for journalism education

In this final section, I will discuss how my analysis could inform journalism education practice.

Firstly, I have highlighted the in-built material injustices of the established power structures data journalism appeals to. Although these are concealed in claims about progress and benefits, journalism educators could consider the implications for their students and the wider community. For example, we should alert students explicitly to the ways in which 'easy-to-use' technology tools channel us to do tasks in a certain way whilst obscuring alternatives. Although 'no coding required' is presented as democratising and empowering, it deskills students conferring power to Big Tech. Learning to code seems burdensome but it potentially liberates students from this dependency, empowering them to seek new ways to explore, share and interpret data.

Secondly, educators can consider who is excluded when we reproduce these discourses in the classroom. Rather than presenting data and technology as neutral artefacts, students need to consider who is not visible in the data, whose interests do technology companies perpetuate and who is not part of the 'everyone' supposedly able to make effective use of open data. Including alternatives to neoliberal, western-centric concepts of journalism and what it considers newsworthy could minimise the exclusionary effects. For example, data skills could be put to the service of issues which matter to minoritised communities.

Thirdly, by identifying four competing, interwoven discourses, my analysis shows there are multiple subject positions made available in data journalism. As educators, we can guide our students to alternative subject positions within this field, drawing on some aspects of these discourses whilst resisting others. Do we position data literacy and technical competency as a means for acquiring competitive advantage in a precarious world or do we configure them as means for students to work on themselves and gain fulfilment?

None of this is to say that data journalism's discourses are bad or harmful but to show that they should not go unexamined (Foucault, 2003). The aim is to highlight the need for educators to continue questioning data journalism's connections to new and powerful apparatus, and how these are channelled into the classroom.

Conclusion

By operationalising Foucault's conceptualisation of discourse in this analysis, I have shown how data journalism produces knowledge about itself and the strategic power plays used by its pioneers to position this as 'good journalism.' The analysis has argued that data journalism draws on four discourses – journalism, technology, enterprise and citizenship – and wraps itself in the power relations embedded in these prestigious discourses. By reproducing themselves in material practices (terminology, physical space, tools), these discourses are intensified and become more 'true.' I argue that data journalism has 'taken' as a recognised meaning because it is so embedded in existing power networks. These are buried under grandiose claims of benefits for all making the power effects "tolerable."

There are limitations to this approach. I have selected a narrow range of pioneer talk but other materials could also contribute – job adverts, teaching materials, online forums. A study of this kind can only ever be partial and so the claims made here are specific to this data set. But it is hoped that this analysis will contribute to an on-going conversation about what Foucault has to offer educators - and data journalism scholarship more widely - by emphasising the ways in which the discursive and the material are inextricably linked to rationalise what it is possible to say. Further research could explore how today's generation of practitioners construct data journalism. To what extent do they reproduce or resist the discourses identified in the expert talk of their predecessors? This could facilitate a fresh understanding of who we are asking our students to become when we teach them data journalism.

Declaration of conflicting interests

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

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Liz Hannaford https://orcid.org/0000-0001-6026-3484

References

Allen DS (2008) The trouble with transparency: the challenge of doing journalism ethics in a surveillance society. *Journalism Studies* 9(3): 323–340.

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

Appelgren E and Nygren G (2014) Data Journalism in Sweden: introducing new methods and genres of journalism into "old" organizations. *Digital Journalism* 2(3): 394–405.

Baack S (2018) Practically engaged: the entanglements between data journalism and civic tech. *Digital Journalism* 6(6): 673–692.

Bazzul J (2015) Tracing "Ethical Subjectivities" in science education: how biology textbooks can frame ethico-political choices for students. *Research in Science Education* 45(1): 23–40.

Berret C and Phillips C (2016) Teaching data and computational journalism. Available at: https://journalism.columbia.edu/system/files/content/teaching_data_and_computational_journalism.pdf (accessed 23 October 2017).

Borger M, Van Hoof A, Costera Meijer I, et al. (2013) Constructing participatory journalism as a scholarly object: a genealogical analysis. *Digital Journalism* 1(1): 117–134.

Borges-Rey E (2016) Unravelling data journalism: a study of data journalism practice in British newsrooms. *Journalism Practice* 10(7): 833–843.

Bounegru L and Gray J (2021) *The Data Journalism Handbook: Towards a Critical Data Practice*. Amsterdam: Amsterdam University Press.

Bradshaw P (2018) Data journalism teaching, fast and slow. *Asia Pacific Media Educator* 28(1): 55–66.

Bucher T (2017) 'Machines don't have instincts': articulating the computational in journalism. *New Media & Society* 19(6): 918–933.

Carlson M (2015) Metajournalistic discourse and the meanings of journalism: definitional control, boundary work, and legitimation. *Communication theory* 26(4): 349–368.

Clegg S (1989) *Frameworks of Power*. London: Sage.

Coddington M (2019) Defining and mapping data journalism and computational journalism: a review of typologies and themes. In: Eldridge SA and Franklin B (eds), *The Routledge Handbook of Developments in Digital Journalism Studies*. Oxon: Routledge, pp. 225–236.

Creech B and Mendelson AL (2015) Imagining the journalist of the future: technological visions of journalism education and newswork. *The Communication Review* 18(2): 142–165.

Creech B and Nadler AM (2018) Post-industrial fog: reconsidering innovation in visions of journalism's future. *Journalism* 19(2): 182–199.

Davies K and Cullen T (2016) Data journalism classes in Australian universities: educators describe progress to date. *Asia Pacific Media Educator* 26(2): 132–147.

De Burgh H (2000) *Investigative Journalism: Context and Practice*. London: Routledge.

De Maeyer J, Libert M, Domingo D, et al. (2015) Waiting for Data Journalism: a qualitative assessment of the anecdotal take-up of data journalism in French-speaking Belgium. *Digital Journalism* 3(3): 432–446.

Deuze M (2006) Global journalism education: a conceptual approach. *Journalism Studies* 7(1): 19–34.

Deuze M and Witschge T (2018) Beyond journalism: theorizing the transformation of journalism. *Journalism* 19(2): 165–181.

Fairclough N (2003) *Analysing Discourse: Textual Analysis for Social Research*. New York; London: Routledge.

Fisher E (2010) *Media and New Capitalism in the Digital Age*. New York: Palgrave Macmillan.

Foucault M (1981) Order of discourse. In: Young R (ed), *Untying the Text: A Post-structuralist Reader*. Boston, MA: Routledge & Kegan Paul, pp. 47–78.

Foucault M (1990) *The Will to Knowledge*. Harmondsworth: Penguin.

Foucault M (2002) *Archaeology of Knowledge*. London: Routledge.

Foucault M (2003) So is it important to think? In: Rabinow P and Rose NS (eds), *The Essential Foucault: Selections from the Essential Works of Foucalt: 1954-1984*. New York: New Press, pp. 170–173.

Gambini L (2019) *A Decade of Working in Data Journalism: What Has Changed?* Available at: https://medium.com/we-are-the-european-journalism-centre/a-decade-of-working-in-data-journalism-what-has-changed-8d950d99935e (accessed 24 April 2019).

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

Green S (2018) When the numbers don't add up: accommodating data journalism in a compact journalism programme. *Asia Pacific Media Educator* 28(1): 78–90.

Gurstein MB (2011) Open data: empowering the empowered or effective data use for everyone? *First Monday* 16(2). doi: 10.5210/fm.v16i2.3316. Available at: https://firstmonday.org/ojs/index.php/fm/article/view/3316/2764 (accessed on 2 February 2019).

Hardy C and Thomas R (2014) Strategy, discourse and practice: the intensification of power: strategy, discourse and practice. *Journal of Management Studies* 51(2): 320–348.

Hepp A and Loosen W (2019) Pioneer journalism: conceptualizing the role of pioneer journalists and pioneer communities in the organizational re-figuration of journalism. *Journalism* 22(3): 577–595.

Heravi BR (2019) 3WS of data journalism education: what, where and who? *Journalism Practice* 3(3): 349–366.

Heravi B, Cassidy K, Davis E, et al. (2021) Preserving data journalism: a systematic literature review. *Journalism Practice* Epub ahead of print DOI: 10.1080/17512786.2021.1903972

Hewett J (2015) Learning to teach data journalism: innovation, influence and constraints. *Journalism* 17(1): 119–137.

Hook D (2001) Discourse, knowledge, materiality, history: foucault and discourse analysis. *Theory & Psychology* 11(4): 521–547.

Howard AB (2014) The art and science of data-driven journalism. Available at: https://academiccommons.columbia.edu/doi/10.7916/D8Q531V1 (accessed 18 May 2017).

Karlsen J and Stavelin E (2014) Computational journalism in Norwegian newsrooms. *Journalism Practice* 8(1): 34–48.

Kreiss D and Brennen JS (2016) Normative models of digital journalism. In: Witschge T and Anderson CW (eds), *The Sage Handbook of Digital Journalism*. London: SAGE, pp. 299–314.

Latour B (1994) On technical mediation — philosophy, sociology, genealogy. *Common Knowledge* 3(2): 29–64.

Lewis NP and Waters S (2017) Data journalism and the challenge of shoe-leather epistemologies. *Digital Journalism* 6(6): 719–736.

Loosen W, Reimer J and De Silva-Schmidt F (2017) Data-driven reporting: an on-going (r)evolution? An analysis of projects nominated for the Data Journalism Awards 2013–2016. *Journalism* 21(9): 1246–1263.

Mayr A (2008) *Language and Power: An Introduction to Institutional Discourse*. London: Continuum.

Mensing D (2010) Rethinking [again] the future of journalism education. *Journalism Studies* 11(4): 511–523.

Ojo A and Heravi B (2018) Patterns in award winning data storytelling: story types, enabling tools and competences. *Digital Journalism* 6(6): 693–718.

Peters MA, Besley AC, Olssen M, et al. (2009) *Governmentality Studies in Education*. Rotterdam: Sense Publishers.

Pilkington E (2015) *Edward Snowden Calls for Global Push to Expand Digital Privacy Laws*. The Guardian. Available at: https://www.theguardian.com/us-news/2015/sep/24/edward-snowden-international-laws-digital-privacy-video (accessed 20 February 2022).

Popkewitz TS and Brennan MP (1998) *Foucault's Challenge: Discourse, Knowledge, and Power in Education*. New York: Teachers College Press.

Powers M (2012) In forms that are familiar and yet-to-be invented" American journalism and the discourse of technologically specific work. *Journal of Communication Inquiry* 36(1): 24–43.

Royal C (2010) The journalist as programmer: A case study of the New York Times interactive news technology department. *International Symposium in Online Journalism*. Austin, TX, USA: The University of Texas at Austin.

Royal C, Bright A, Pellizzaro K, et al. (2020) Product management in journalism and Academia. *Journalism & Mass Communication Quarterly* 97(3): 597–616.

Splendore S, Di Salvo P, Eberwein T, et al. (2016) Educational strategies in data journalism: a comparative study of six European countries. *Journalism* 17(1): 138–152.

Springer RA and Clinton ME (2015) Doing Foucault: inquiring into nursing knowledge with Foucauldian discourse analysis. *Nursing Philosophy* 16(2): 87–97.

Stalph F and Borges-Rey E (2018) Data Journalism Sustainability: an outlook on the future of data-driven reporting. *Digital Journalism* 6(8): 1078–1089.

Treadwell G, Ross T, Lee A, et al. (2016) A numbers game two case studies in teaching data journalism. *Journalism & Mass Communication Educator* 71(3): 297–308.

Van Dijck J (2014) Datafication, dataism and dataveillance: big data between scientific paradigm and ideology. *Surveillance & Society* 12(2): 197–208.

Wetherell M, Taylor S and Yates SJ (2001a) *Discourse as Data: A Guide for Analysis*. London: Sage.

Witschge T, Anderson CW, Domingo D, et al. (2019) Dealing with the mess (we made): unraveling hybridity, normativity, and complexity in journalism studies. *Journalism* 20(5): 651–659.

Yang F and Du YR (2016) Storytelling in the age of big data: Hong Kong students' readiness and attitude towards data journalism. *Asia Pacific Media Educator* 26(2): 148–162.

Zelizer B (2005) Definitions of Journalism. In: Overholser G and Jamieson KH (eds), *Institutions of American Democracy: The Press*. New York: Oxford University Press, pp. 66–80.

Appendix

Key speaker	Code	Source	Platform	Date
Simon Rogers	A1	http://www.niemanlab.org/2010/08/how-the-guardian-is-pioneering-data-journalism-with-free-tools/https://vimeo.com/13790996	Nieman lab	August 2010
	A2	https://www.youtube.com/watch?time_continue=3&v=h2zbvmXskSE	TEDx talks/*The Guardian*	December 2012
	A3	https://www.youtube.com/watch?v=yt8dhOXv74s&list=UUJ1NGNrj5b_x789iOLyFr2g&index=1	Global editors' network	March 2013
	A4	https://vimeo.com/90656079	European journalism centre	May 2014
Aron Pilhofer	B1	https://medium.com/tow-center/aron-pilhofer-on-data-journalism-culture-and-going-digital-409c7c43c0f8	Tow center for digital journalism	2014
	B2	https://www.youtube.com/watch?v=GN4mC-o5WzY	Centre for investigative journalism	July 2017
	B3	https://www.youtube.com/watch?v=OScmB1f1jM8	International journalism festival	April 2018
Paul Bradshaw	C1	https://www.aendrew.com/content/evolving-newsroom-qa-paul-bradshaw	Aendrew.com	May 2012
	C2	https://vimeo.com/90960252	European journalism centre	May 2014
	C3	https://www.youtube.com/watch?v=5od0WIg72BA	Nordic data journalism conference	May 2016
Adrian Holovaty	D1	https://www.theguardian.com/media/audio/2008/jun/12/adrian.holovaty	*The Guardian* "future of journalism" conference	June 2008
	D2	https://www.theverge.com/2012/8/16/3245325/5-minutes-on-the-verge-with-adrian-holovaty-founder-of-everyblock	*The Verge*	August 2012
Brian Boyer	E1	https://itsalljournalism.com/nprs-apps-editor-brian-boyer-turns-data-into-stories/	*It's all journalism*	July 2013
	E2	https://source.opennews.org/articles/boyer-interview/	*Source* (part of open news)	November 2013
Michelle Minkoff	F1	http://radar.oreilly.com/2012/03/profile-of-the-data-journalist-3.html	O'Reilly.com	March 2012

(continued)

(continued)

Key speaker	Code	Source	Platform	Date
Jacqui Maher	G1	https://www.youtube.com/watch?v=864pOLNWsZo&feature=youtu.be	International journalism festival	April 2015
Alberto Cairo	H1	https://vimeo.com/92064785	European journalism centre	May 2014
	H2	http://www.niemanlab.org/2014/12/visualization-goes-mainstream/	Nieman lab	December 2014
Nicolas Kayser-Bril	I1	https://www.youtube.com/watch?v=We7WfjUVhLs	European journalism centre	November 2010
	I2	https://vimeo.com/92042656	European journalism centre	May 2014
	I3	https://www.youtube.com/watch?time_continue=5&v=zv8Fs8snoUE	International journalism festival	April 2015
	I4	https://www.youtube.com/watch?time_continue=18&v=aXgrC4A16Mo	International journalism festival	April 2015
Megan Lucero	J1	http://www.interhacktives.com/2015/12/21/megan-lucero-interview/	Interhacktive, city university, london	December 2015
	J2	https://www.youtube.com/watch?v=Smyd5QxXyjI	International journalism festival	April 2016
	J3	http://datadrivenjournalism.net/news_and_analysis/getting_with_the_times_lessons_from_the_papers_data_journalism_editor	Data-driven journalism	December 2016
Sisi Wei	K1	https://medium.com/tow-center/data-skills-make-you-a-better-journalist-says-propublicas-sisi-wei-fb16e7370f4	Tow center for digital journalism	April 2014

Author Biography

Liz Hannaford lectures in multimedia journalism at Manchester Metropolitan University. Her research interests are data journalism, discourse analysis, pedagogy and subjectivities. Prior to becoming a lecturer, she worked for 20 years as a broadcast journalist at BBC World Service.