DiVA
http://www.diva-portal.org

This is the published version of a paper published in *Journalism Practice*.

Citation for the original published paper (version of record):

Morini, F. (2025)
Data Journalism as "Terra Incognita": Newcomers' Tensions in Shifting Towards Data Journalism Epistemology
Journalism Practice, 19(1): 76-92
https://doi.org/10.1080/17512786.2023.2185656

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:sh:diva-51234

JOURNALISM PRACTICE
https://doi.org/10.1080/17512786.2023.2185656

Routledge
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Data Journalism as "Terra Incognita": Newcomers' Tensions in Shifting Towards Data Journalism Epistemology

Francesca Morini [a,b]

[a]Södertörn University, Sweden; [b]University of Applied Sciences Potsdam, Germany

ABSTRACT

This article investigates data journalism epistemology through Michel Foucault's definition of power. The growing demand for data-savvy reporters with computational skills has been proven to shift the newsrooms' culture in media companies across the globe. Previous research has documented journalists' shift towards a data-centred epistemology and the increasingly important role of computation and data-driven practices in newsrooms. By focusing on inexperienced journalists as they mobilise data journalism for the first time, this research openly discusses its epistemology as a form of Foucauldian power: which epistemological tensions – individual or structural – emerge among newcomers when reporting extensively with data? The article presents the case of a newly formed data journalism team at taz, a German cooperative media company. It is based on qualitative data collected through action research, namely six in-depth interviews, auto-ethnographic notes and self-reflective inquiries. Data are framed through Foucault's concepts of power and will to truth. Results show the discord between data journalism epistemology and its practical implementation. Despite prioritising democratic access to data, technologies and non-hierarchical forms of cooperation, data journalism remains largely dependent on traditional methods and culture.

ARTICLE HISTORY
Received 7 August 2022
Accepted 19 February 2023

KEYWORDS
Data journalism; data journalism epistemology; transdisciplinarity; action research; innovation; cooperative journalism; practice-based research

Introduction

The rapid growth in popularity of data journalism is shaping traditional newsrooms' culture and workforce (Kosterich and Weber 2018). Journalists are increasingly asked to work with data and new technologies (ibid.), and non-traditional actors or "in-betweeners" of journalism are becoming central to the reporting work, contributing new knowledge, skills, and production patterns (Ahva 2017). Experienced journalists are increasingly confronted with computer enhanced methodologies and how to embed data journalism in their everyday work. This transition is far from being linear and uncomplicated, especially in newsrooms with limited resources (Stalph and Borges-Rey 2018): journalists implement a variety of strategies to upskill and deal with new tasks involving advanced

CONTACT Francesca Morini francesca.morini@sh.se
This article has been corrected with minor changes. These changes do not impact the academic content of the article.

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

technical knowledge. This transition marks an epistemological shift for journalists changing the way in which they produce knowledge and epistemic validity, in other words how they justify the righteousness of their claims on the empirical world. Data-based methods and practices change the way in which journalists make sense of information and connect evidence of certain events in order to craft trustable news. The epistemology of data journalism – namely the origins and cultural affordances of this practice – is an implicit bundle of rules that shapes how the practice is looked at and accessed. Previous research has largely focused on the epistemology of data journalism and how it influences newsrooms' culture and journalists' approach to reporting. Several contributions have highlighted the differences and similarities of data journalism in comparison with other forms of data enhanced journalism like Computer Assisted Reporting or wide-spread genres such as investigative journalism (e.g., Parasie and Dagiral 2013; Coddington 2015; Parasie 2015). Other scholars focused on investigating the cultural, methodological and conceptual imports to journalism from a variety of fields, including computer science (Gynnild 2014; Lewis and Usher 2014), design (Usher 2020), and data studies (Lewis and Westlund 2014). Borges-Rey (2020) focused on how experienced data journalists switch between different epistemological approaches when working on data journalism projects. Still, little attention has been paid to the strategies that newcomers of data journalism adopt when working with data stories to produce reliable and sound reporting. With this study, I analyse how newcomers make sense of working with data and their struggle in adhering to its epistemology to produce truthful and reliable claims on the empirical world. In this context, the term newcomer is used to identify journalists skilled in traditional journalistic techniques that attempt at accessing data journalism and stabilise it as a potential recurring practice in their everyday workflow. Arguably, this switch from a traditional to a data oriented epistemology is a source of tension – both individual and structural – for news professionals. Resolving such tensions has direct consequences on the newsroom's culture and changes the way in which data journalism is approached from a broader perspective. Observing this process of adaptation is relevant to understanding the influence of data journalism epistemology on newsrooms' culture and the conditions for data journalism to be systematically implemented. Therefore, the following research question (RQ) will be addressed:

(1) *Which epistemological tensions – individual or structural – emerge among a team of newcomers when reporting extensively with data?*

To this extent, I will approach data journalism epistemology using Michel Foucault (1971) definitions of "power" and "will to truth". Power is the sum of all constraints, boundaries, and instances that constitute a specific domain of society (Gershon, Eick, and Card 1998). Will to truth designates the ensemble of the material, technical, and instrumental investments of knowledge necessary to access and mobilise power (Foucault 1971). By looking at data journalism epistemology as a form of power, I will report on journalists' will to truth as they gain epistemic validity in performing data journalism. Since neither power or will to truth can be accessed from outside their domain (ibid.), I will use action research (Stringer 2013) as a method to unravel epistemological tensions in the making of a data story. To this extent, I joined a newly formed data journalism team at *die tageszeitung* – a German cooperatively-owned newspaper – and

developed a one year long journalistic investigation. By providing practical support to the team in developing static and interactive data visualisations (Gershon, Eick, and Card 1998) for the purpose of storytelling (Segel and Heer 2010), one data story was produced on the topic of development funds. During the project, I collected data through six in-depth interviews, field annotations, and auto-ethnographic records. As a result, three unique tensions emerging from working together with a newly formed team will be reported and analysed to explore how data journalism is produced, legitimised, and stabilised as a recurring practice.

Data Journalism in Germany and at Die Tageszeitung

The findings presented in this study originate from one individual case study. Collecting data from unique case studies allows for direct observation of real-life events where phenomenon and context are not readily separable (Yin 2004). For one year, I joined a journalistic team at *die tageszeitung* (shortened taz), a daily German newspaper based in Berlin. Our goal was to collect and analyse data for a journalistic investigation. The initial premises relied upon the intuition that the consistent funding for water development projects granted by Germany to other nations could reveal something about the country's diplomatic and political influence. During the project various datasets were analysed to uncover ties between countries, governmental institutions, and semi-public consultancy agencies across Germany. Several revelations were made: development money is allocated not solely on the basis of necessity as geopolitics is often at stakes, migration crises play a fundamental role in the soaring costs of development projects and, it is not possible to transparently track which companies got hold of the money once they are disbursed by the ministry. Taz is a cooperative-owned media company established in 1978 (Groll 2012*)*. Media cooperatives are companies characterised by a unique business model: employees and subscribers directly own shares of the newspaper (Siapera and Papadopoulou 2016). Cooperatives base their work on rigorous journalistic standards, democratic participation, voluntary and open membership, and self-responsibility (ibid.). They are typically more financially resilient than traditional media outlets, although they are often constrained by limited resources (Siapera and Papadopoulou 2016). Previous research within and beyond the German context has shown how data journalism is still difficult to implement for less-resourced newsrooms (Stalph and Borges-Rey 2018; Stalph, Hahn, and Liewehr 2022), whereas seeked after by reporters and editorial teams alike. At taz, journalists are working towards implementing data journalism as a recursive and cohesive practice, while dealing with restricted resources. The data journalism and tech scene in Berlin are particularly active, especially within legacy newsrooms (Baack 2018), whereas at taz the process of changing individual and newsroom culture to accommodate data journalism is still in progress and allows for fruitful observation of emerging tensions in implementing it. To this extent, taz poses the ideal context to study journalists' epistemological transition towards data-driven reporting.

Data Journalism and Its Epistemology

Data journalism is a form of storytelling, where traditional journalistic methods are mixed with programming, data science and visualisation and the ethos of open-source culture

has a central role (Borges-Rey 2016; Appelgren and Nygren 2014; Coddington 2015). Given its aspirations, data journalism sits at the intersection between technological innovation and journalistic tradition (Gynnild 2014). For practitioners, data are considered the primary source and news material (ibid.). In using data as a source and approaching them from a clear news a gle, data journalists closely link themselves to professional views of journalism (Coddington 2015) while sporting a technology-oriented mindset (Gynnild 2014). Several scholars have focused on observing how journalists consider, organise, and communicate data in order to back-up claims resulting from their reporting activities. Looking at data journalism through its epistemology – its boundaries and how journalists build epistemic validity by leveraging its techniques (Parasie and Dagiral 2013; Ekström, Ramsälv, and Westlund 2022; Borges-Rey 2020) – is useful to determine its influence on audiences and other news professionals alike. While the role of data clearly permeates the discourse on data journalism epistemology (e.g., Lewis and Westlund 2014), there are also other key aspects connected to this profession's origin and what it entails for journalists exercising it as their everyday practice. Technology, described as a driver for innovation, and transdisciplinary cooperation, which support the spread of open-source culture within data journalism teams (Appelgren, Linden, and van Dalen 2019), are also considered central. To summarise data, technology, and transdisciplinarity are key to the epistemology of data journalism and have been discussed to an extent both by practitioners and scholars.

Data is an Objective and Transparent Source of Truth

Data is here defined as information in an electronic form, capable of being stored, processed, and accessed by a computer (McIntosh 2022). Every piece of information systematically stored and recorded can be considered data. Journalists with advanced technical skills become able to use data and not pre-analysed or aggregated records, namely "raw data" (Appelgren and Nygren 2014). This ability is considered as particularly substantial and contributes to data journalists status within the newsroom (ibid.). Data journalists themselves remark how their often self-taught knowledge of statistics, coding, and computer science is undeniably better than their colleagues' (Haim 2022). Through data analysis and visualisation data journalists become able to build stories and therefore capable at producing valid claims about the empirical world (Usher 2020). In this sense, the production of data driven artefacts is a form of authority over data (ibid.). Similarly, data savvy journalists claim to be capable of questioning data as they would do with traditional sources (Parasie and Dagiral 2013). Their work is seen as more systematic, pioneering, and trustworthy (Zamith 2019) – as they tap into data as a source to unveil patterned evidence and hidden truths (Haim 2022).

Data Journalism is Technology-Driven and Democratically Accessible

Data journalists' ethos is largely based on their do-it-yourself approach that leads them to learn all the necessary methodologies to carry out journalistic work even when moving in uncharted territories (Hepp and Loosen 2021). Here lies the central aspect that distinguishes Computer Assisted Reporting (CAR) from data journalism. The former relies on the idea of data as unstructured knowledge and on developing new methods and algorithms to support journalists in making sense of this information (Coddington

2015). Its outcomes seldomly include the publication of data to the reader benefit. The latter has a strong narrative focus (ibid.). Data journalists are committed to finding stories within data and developing novel formats to tell these stories to their readers (Gynnild 2014; Heravi and Lorenz 2020): data, and the technology involved in extracting them, become key components of the story. Technological tools are part of data journalists' everyday routines as a strategic shortcut to process and work with large datasets (Appelgren and Nygren 2014). Meanwhile, technology is often described as an enabler for data journalism within newsrooms and as accessible to everyone, not only specialists (Gynnild 2014; Bounegru and Gray 2021). Every journalist can experiment with this practice as plenty of resources are made accessible online (Bounegru and Gray 2021). Reporters often develop new tools to obtain and analyse data or customise existing ones (Gynnild 2014). Their active engagement in these communities is a way to build authority and social capital among other professionals within the field.

Data Journalists are Empowered by Transdisciplinary Exchange

Democratic access to resources and technologies often transcends geographical, linguistic, and organisational levels (Gynnild 2014). Data journalists are often looking for connection and cooperation outside of their team or newsroom (e.g., Porlezza and Splendore 2019; Appelgren, Linden, and van Dalen 2019). Collaboration makes it possible for journalists to move away from competitive environments and establish a network of like minded professionals (de-Lima-Santos 2022). The transdisciplinary nature of the practice – meaning that projects need to be approached at the intersection between disciplines (Szilárdi and Tőke 2019) – is praised as a key factor in producing cutting edge reporting. Teams are increasingly transdisciplinary, welcoming programmers, designers, and scientists (Ahva 2017) motivated to interlock by common values and ambitions (Baack 2018). Even if tensions among practitioners emerge, they are often seen as necessary for the creative process (Parasie 2015). Transdisciplinarity is not limited to exchange with external actors, journalists themselves are found to change between different epistemological approaches when working with data (Borges-Rey 2020). Computational and journalistic mind sets are interchangeable; as a result transdisciplinarity gains centrality in crafting ground-breaking editorial products (ibid.).

These three aspects have been variably described as foundational to data journalism epistemology. They grant data journalists agency and credibility among other professionals and the general audience. Nonetheless, mobilising such elements is a costly and complex enterprise for journalists (Parasie 2015) leading to several epistemological tensions. Newcomers of data journalism struggle to trust data, utilise technology, and work functionally with team members (ibid.). On the one hand, reporting on such tensions from the point of view of newcomers helps in understanding how data journalists build epistemic validity. On the other hand, it supports the observation of newsroom's culture is influenced and changed over time as data journalism gains relevance.

Approach

So far, journalism scholars have largely focused on understanding how experienced data journalists validate their claim and therefore construct a certain epistemology of their

practice (e.g.,: Appelgren and Salaverría 2019; de-Lima-Santos and Mesquita 2021; Mutsvairo 2019). Resulting epistemological frameworks are attempts at mapping data journalism through its constraints and relations. Data journalism is approached as a way of constituting knowledge. In this paper, I will rather look at data journalism epistemology as a form of power: by tapping into it, journalists become capable of claiming epistemic validity on real world events through its strategic use. According to Foucault (1978), power must be understood as the attributes to a complex strategic situation in a particular society. Power is a sum of all constraints, boundaries, and instances that constitute and influence a specific domain (ibid.). It is not gained or externally imposed, but permeates all existing relations, subjects, and objects (ibid.). Power is exercised and manifested through the will to truth: "the range of objects to be known, of the functions and positions of the knowing subject, of the material, technical, and instrumental investments of knowledge" (Foucault 1971, p. 51). Power is not inherently good or evil, but a moving substrate of relations without centrality that require constant negotiation (Foucault 1978, p. 93). In this sense, data journalism epistemology can be considered a form of power, its constraints being constantly redefined and mobilised by its practitioners. Here the role of data, technology, and transdisciplinarity comes into play: as key constraints, they also characterise journalists' "will to truth", hence their way of constituting reality. These three aspects also play a fundamental role in shaping data journalism as a truthful, transparent, and reliable practice. In this paper, the main goal is to observe data journalism epistemology as a form of Foucauldian power by looking at newcomers' struggle to access data journalism epistemology and gain control over it. By using an action research approach (Stringer 2013), research work was conducted in a participatory fashion, by blending practice and academic investigation (ibid.). Action research is a fundamentally qualitative approach that mixes participatory and reflexive methods, providing the means to practically engage with relevant stakeholders within their context (Stringer 2013; Hearn et al. 2009). It is an approach useful to develop new knowledge through new understanding of situations (Hearn et al. 2009) and to foster proactive change. By allowing proximity between practitioners and scholars, action research is a powerful method to observe rapidly changing and converging professional environments (Wagemans and Witschge 2019) as well as to enable and foster innovation (Stringer 2013; Grubenmann 2016). In this study, I will firsthand experience how these three key aspects of data journalism – data, technology, and transdisciplinarity – shape reporting and how newcomers deal with them. To this extent, action research is a fruitful approach to study epistemology as a form of power, allowing for observation of real-life events together with their context and actors. Moreover, the intersection between journalism and academia is particularly fit in this context as it fosters for transdisciplinary cooperation, creating the condition for data journalism to happen.

Materials and Methods

Action research has been previously employed in journalism studies for different scopes and contexts (e.g., Appelgren and Nygren 2014; Grubenmann 2016; Wagemans and Witschge 2019; Picha Edwardsson 2022); its phenomenological, interpretative, and hermeneutic premises allow for researchers to gain direct experience from specific situations (Stringer 2013). In this case, action research was employed in combination with reflexive

practices, such as semi-structured interviews, ethnography (Eriksson, Henttonen, and Meriläinen 2012; Naidoo 2012), and self-reflective note taking (Marshall 2001). The project took place between December 2020 and October 2021 and resulted in an extensive process of collaboration between journalists, editors, academics, and programmers. The cooperation was initiated by a dialogue between editors at taz and me, as a representative of my research institution. The goal was to publish a data story on the topic of German development politics. More generally, the idea was to develop a pipeline for data journalism, comprising technical and methodological infrastructure to facilitate future journalistic work. To this scope, a data journalism team was formed, including one reporter, one developer, and one digital products' specialist in charge of editorial decisions. I joined as developer and designer, to devise and deploy charts and visual elements, complementing the skillset of the team. Due to the ongoing COVID-19 pandemic, all interviews and meetings were conducted remotely. The research resulted in one longform article – namely a digital story with multimedia assets spanning at length over one topic (Dowling and Vogan 2015) – published on taz's online platform and across social media. Additionally, one podcast episode was also produced and released online one month after the publication.

All key actors involved in the production team were interviewed at the start, middle and end of the project. To ensure ethical rigorousness, the research was submitted for ethical review. All participants were asked for consent before being interviewed as well as prior to note taking. Interviews were conducted with the full consent of people involved. Participants were informed about the length and goal of interviews and were subsequently asked permission to use key extracts for this manuscript. Moreover names and references to colleagues and other people involved were removed in the data collection phase. Six interviews in total were conducted, all of them in English. Initial interviews were structured to assess participants' perceptions about the project and position within the team, whereas final ones aimed at recording new and changing impressions and validating previously made claims (Lincoln and Guba 1985). In final interviews, questions were constructed by scanning ethnographic notes and other materials resulting from the project and teasing out interesting elements and observations made by the researcher during the cooperation period. Ethnographic observations and self-reflexive inquiries were recorded in the form of field notes. This structure for note-taking was devised to highlight aspects of the communication tied to epistemological tensions. Following Marshall's (2001, 433) paradigm for self-reflective inquiry, it is fundamental to move between subjective and objective information to record both personal observations and the surrounding context. According to this method, it is necessary to rigorously observe the researcher's own thoughts and mental activities, such as alternance between thinking and doing, figures of speech and real time annotations during meetings that account for feelings and first-hand impressions (ibid.). Interactions with team partners and context have been recorded through ethnographic notes, oriented at recording material and social relations (ibid., p. 434; Eriksson, Henttonen, and Meriläinen 2012). The material was analysed against the three aspects of data journalism epistemology previously introduced: data, technology, and transdisciplinarity. Initial interviews and ethnographic notes were further analysed according to three textual dimensions: what is recorded, who is central in the recordings, and how events are described (Eriksson, Henttonen, and Meriläinen 2012, 14–16). From this first round of

analysis, some recurring topics – related to Foucault's notions of power and will to truth – were leveraged following a deductive model, to analyse the final round of interviews with project's members.

Results

Findings highlight the unfolding of tensions connected to the previously introduced aspects: (T1) the difficulty in trusting data and working with them productively, (T2) the challenge of managing restricted financial and professional resources, and (T3) the perceived disparity among team members regarding project's ownership. These tensions emerged for some of the key actors individually as well as among team members and have been dealt differently from time to time. In the upcoming sections, each tension is described in detail, discussed, and contextualised against the initial research question.

T1 – Data as Terra Incognita

The team started with a large database comprising all transactions between states involving money dispatched to water development. This source is public and has been designed to report all funded development projects since 1970. The story was initially planned entirely around this resource: team members were set to extract information on which institutions are responsible for managing development money and which countries benefit the most from these funds. When asked to the reporter how she would approach data, she initially had a very clear idea of how to mobilise them within her story: "You might have some idea [about the topic] but then the data would tell you if it's not correct or farfetched." (reporter, interview, December 15, 2020) This approach resonates with the reporter's traditional strategy for writing stories. She started from the assumption that to gather information you need to ask questions to relevant actors:

> "I would try to reach out to [who I need to talk to] and ask them for comments […]. You need to gather all the information relevant: twitter updates, last minute events, press agencies releases. […] You try to get authorities to talk to you, who's important? Who are the big actors?" (reporter, interview, December 15, 2020)

The connection between data and traditional journalistic sources becomes explicit: the reporter intends to "ask" questions to data and they would "tell" her if ideas and hypotheses are correct. During the course of the project this approach shifts consistently towards a more faceted view on how to treat data as a central source: "[…] when you get deep into [data], you also pose other questions on how to analyse them. […] The more questions I had the harder it was to say "*this is the stop, we don't go further.*"" (reporter, interview, November 26, 2021). If at the beginning data held the same status as a normal source, by the end of the project the reporter realised how daunting it was to validate intuitions and ideas through them. Data started to be considered more broadly as research material that needs to be externally validated. A major tension emerged: data did not fit coherently with her model of individual sense-making and with her initial expectations. During the project the reporter enacted a series of coping strategies in order to trust the information at hand: when confronted with technical obstacles, she resumed very quickly to more traditional journalistic techniques (short interviews, press releases, documentation, etc.) and

leveraged on colleagues' support in making sense of data. These strategies intended to address a very big concern of hers: validating her claims against the data without being left uncertain. The reporter's temporary inability to trust the information at hand made her want to go back in control through what she knew as practical and effective methods of inquiry. This tension is a major pivot point for her as it is cause for friction, it is so difficult to emotionally cope with that it actively turns into an obstacle to work with data:

> "[…] I think that this thing [the data research and analysis process] was a terra incognita where I was blind, sitting in front of all these data, always going back and forth asking "Is that in here? Is that there?"" (reporter, interview, November 26, 2021).

This feeling of being lost and powerless in front of data affected all team members variably. One account of this tension emerged from ethnographic notes: the difficulty in coordinating data work. At the beginning of the project, my first task was to gather all data extracted from the analyst assigned to the project and make it available for the reporter in a more friendly way, through a series of interactive documents. These reports have been largely overlooked, as the reporter preferred to explore data threads independently. This inability of committing to a common strategy prompted several switches between datasets through the project, which caused the journalist to question her reporting abilities and challenged my position within the team. This tension is reflected in the increasing frustration both from both sides – mine and the reporters' – in balancing these two different approaches for data exploration. The tension itself did not directly reflect on the quality and reception of the final product. According to the reporter herself and the rest of the team, the story was well received and commented both outside and inside taz, with extensive praise by the public and other professionals.

T2 – Managing Innovation with Restricted Resources

Without the support of a dedicated data team with renowned ability to work on data journalism projects, the reporter found herself struggling, unable to fruitfully mobilise her training. As reported by the editor in charge of the project, this lack of professional figures specialised in data journalism within the newsroom is largely due to taz business model: the lack of economic resources makes it difficult for them to hire programmers and data analysts with journalistic experience. Restricted capital is a big obstacle to taz's internal innovation, as professionals are left with little support and have to rely on their own knowledge and skills. In this particular context, the editor in charge of the project was the only figure with a journalistic background mixed with technical skills, and therefore capable of providing infrastructure and support to all members in the team. With her competencies, the editor took particular care in shepherding the reporter through the project. As the reporter herself commented: "I still feel like I don't have the competences I need to do that, and that was also why I was so happy to have the editor on the team. […] I feel like there is a lot of translating to be done, you need someone you can talk to." (reporter, interview, November 26, 2021).

The editor approached the undergoing data journalism project as a chance for innovating newsroom culture, often at the expense of technological innovation itself. Her approach was clearly geared towards two main goals. First, supporting the reporter in using the technological tools necessary to extract and make sense of data. Secondly,

devising a sustainable project pipeline structurally compatible with taz's technological infrastructures and potentially replicable by other journalists. In this context, due to limited resources, the definition of "innovation" is at stake. Again, this proved to be a point of friction between them – journalists and newsroom workers – and me as a technologist. For the editor this concept was tied to sustainability: "It happened in the past that we had money […] and [we] just produced a website, very smooth and interesting. If you look at it from a result perspective […] it makes sense. But […] from a sustainability perspective there are so many disadvantages. The project can't be included in our CMS, it can't be found, and it's not good for SEO" (editor and product specialist, interview, November 3, 2021). For me, innovation remained tied to potentially introducing new formats or ways of representing data. This tension emerges clearly from ethnographic and self-reflexive notes. At times, I describe the project as "too slow" or "fuzzy" and express concerns in regards to the timeline. On several occasions I point out the editor's little interest in regards to visualisation and the limited feedback received, as the focus remained extensively on the journalist's work. This mismatch often led to an overarching feeling of dissatisfaction from my side, which resulted in a strong perceived disparity of ownership.

T3 – Disparity of Ownership

In relation to the previous tension, the overall setup of the project has played its part in influencing team members' expectations and behaviours, especially in regards to the extension and boundaries of our cooperation. Being focused on design, my initial contribution to the project was to prepare multimedia artefacts and to extensively work on original visual formats to report with data. This agreement set the boundaries of our cooperation on equal grounds, where I would have had the freedom to propose external work (e.g., articles, visualisations and story pitches) for their platform. These expectations were systematically dismissed over the course of the project as editorial decisions became often the focus for discussion. Despite the transdisciplinary frame, technical work was often secondary to editorial tasks. This mismatch between the editorial and technical side, where journalistic work was often prioritised over any other type of contribution, prevented me from entirely feeling part of the team and editorial decisions. For instance, rather than running independent analysis on the content, I was asked on several occasions to provide simple visualisation or data cleaning tasks. The lack of shared ownership in the project influenced the way in which I approached assignments, by continuously adjusting my working pipeline and trying to understand how my work can be perceived as useful from other team members. In this sense, my contribution became increasingly service oriented with little agency over the story and its content. Not feeling confident in proposing and motivating ideas to further proceed with data work and story editing had the effect of making my work less effective while at the same time made me question my position and role in the project.

While initially this feeling of disparity emerged in relation to my position within the project, it became evident during the cooperation that other team members were also affected by it. Hence, this disparity is not reflected exclusively with outside partners, but also internally between colleagues at taz. For instance, the developer involved in the project has worked almost exclusively on data extraction and analysis. He did not

perceive himself as part of the project, if not as a consultant and supporter. He claimed to have little or no experience in journalism. His main occupation as taz was normally not directly related to newsroom work, but he lent his experience to colleagues.

> "Journalists are frustrated that they don't get data from the [official] sources. Most of the time journalists need numbers, I know data but I wouldn't know how to write an article. […] I knew the reporter needed help and I stepped in." (programmer, interview, January 31, 2021).

This unbalance makes the discussion around data and findings uneven. Whereas the journalistic perspective was always strongly vouched for, technical aspects such as particular investigation techniques or representation strategies for data were always briefly discussed and never prioritised. The programmer himself implied that his role was oftentimes secondary to journalists' will: "I point some aspects [of data] out. Other people decide what is interesting or not. Oftentimes we are on the same page. […] Sometimes they just ignore it." (programmer, interview, January 31, 2021). Both him and myself were constantly asked to negotiate our position within the project. This setting, while extremely efficient from a production perspective, limited the team work effort and experimental cross-pollination of different practices, as reporting, design, and technical developments remained rigidly separated within the project.

Summary of Results

Findings presented above highlight the challenges of carrying out data journalism at taz, according to three different dimensions: data, access to technology, and transdisciplinarity. Each of these aspects is mobilised during the project and is a source of tension. The first tension (T1) concerns primarily the journalist involved on the project as she struggles to switch back and forth between a traditional and data-oriented epistemology. The second aspect (T2) concerns mainly the lack of a solid infrastructure and network of professionals within the newsroom. To this extent, the editor in charge is responsible for managing team work and pushing for some technical boundaries to be maintained, influencing the final results. The third clash (T3) revolves around the uneven treatment reserved to non journalists within the team. As a result not every team member detains full ownership of the project leading to lack of trust. By observing the unfolding of these tensions, it is possible to look at data journalism epistemology as an issue of power: data, access to technology, and transdisciplinarity need to be understood, made actionable, and taken in by journalists and practitioners in their attempt at tapping into data journalism epistemology. They constitute the key relations at the core of data journalism as they are interconnected forces ordering how the practice has to be carried out. By learning to tap into this set of rules, data journalists gain epistemic authority in producing knowledge through data. Who is able to access and understand data, to leverage the correct methodologies and proper tools, and to communicate findings across professional fields and audiences is also able to access data journalism epistemology.

Discussion

In light of these findings, epistemology can be understood as a set of rules and constraints to generate knowledge, a power strategy that establishes journalistic work made with

data as impartial, reliable, transparent, and accountable. Data journalists themselves actively work to preserve their position of power by presenting their work as objective, trustworthy, democratically accessible and cross-pollinated by other practices. They are seen as capable of discovering hidden truth behind data through a plethora of unique technical skills that fit in the process seamlessly. This holistic image of data journalism contributes to its popularisation and influences how newcomers approach its epistemology. The struggles emerging from working on a project together with reporters owning limited experience in the field reveals how data journalism epistemology is far more complex and up for grabs. Journalists have to continuously move back and forth from traditional and computational methods (Borges-Rey 2020). They change data sources and methodologies if they don't find them to fit their narrative. They are able to grasp a certain degree of truth from data, but they often trust their instincts and colleagues in the process. Ultimately, they struggle to reconcile different perspectives and often prioritise journalistic methodologies and outcomes over others.

Data as a Form of Power

Data journalists' work largely consists in making data transparent to readers, offering them a glimpse of truth hidden within their patterns. In Foucauldian terms, data savviness is to be seen as a tactic to achieve power (Foucault 1975, 95), as it allows the production of reliable and objective knowledge over reality. Knowing how to work with data grants them a semblance of authority. This observation resonates with what has been found by Usher (2020): journalists use data and visualisations to establish epistemic credibility. Similarly, Appelgren and Nygren (2014) discuss data journalism as a way for reporters to gain influence and status within the newsroom. As a newcomer to this tactic, the reporter clearly experienced difficulties as she tried to work with data. She was constantly confronted with the idea of data journalism as objective, data-driven, and systematic: data cannot lie, if approached with the correct techniques. However, when working with statistics or visualisation techniques, she was often producing limited and superficial findings that could be hardly utilised for her story. What became clear as she worked through the project was that data themselves are neither meaningful nor transparent sources of truth. As also stated by Parasie and Dagiral (2013) in their work, sense-making and investigative abilities tied to traditional journalistic practices remain central in organising and contextualising information. To this extent, reporting with data is not exclusively built on journalists' technical skills and background. It requires in-depth knowledge of journalistic methods and standards. Whereas data is often presented as the key object to be known in order to unveil the truth, its function is only secondary to journalists' subjectivity, partially challenging the image offered by Zamith (2019) of data journalism as an inherently trustworthy and systematic practice. These observations allow us to frame data as a constitutive element of data journalism influence. By presenting data as truthful and their work as objective, data journalists reinforce the idea of themselves as outsiders of journalism with increasing authority, as also discussed by Haim (2022). This has potential ramifications both on newcomers, as they struggle to merge different epistemological approaches and tend to prioritise data work as the primary source for truth, and on newsroom culture, as data journalism might remain relatively restricted to smaller professional groups.

Technology and Democratic Access

As noted by the reporter involved in our project, for a beginner, data journalism feels like a magician's work. During our cooperation, the editor and reporter were often taking up a considerable amount of time to discuss how to approach data and to review which tools could be used to produce reliable findings. These discussions never concerned the possibility of experimenting with innovative formats or technologies, but were heavily focused on existing tools and their adoption. Even if resources, techniques, and methodologies for data journalism appear to be democratically accessible to everyone (Coddington 2015; Bounegru and Gray 2021), this condition is not alone sufficient for data journalism to happen. As the key element of innovation is a skilled workforce (Fink and Anderson 2015), access to technology remains marginal if not coupled with a renewed and proactive newsroom culture. If experienced data journalists often consider technology as a shortcut to produce their work (Appelgren and Nygren 2014), editors and project managers should start to see them as a form of disciplinary power. According to Foucault (1975): a way to train newcomers in working with data and gain epistemic authority over them. Discipline is a technique of power to combine different segments of a practice to create a permanent economy (ibid.). This observation potentially challenges the idea of data journalism as a democratically accessible practice that can be easily adopted invariably by all reporters: within newsrooms a systematic approach to technology is key for newcomers' transition to data journalism epistemology. This idea also reinforces what has been previously discussed by Gynnild (2014) in regards to technological innovation in the newsroom. Creating and maintaining new tools is not the only aspect connected to technology, it is also coupled with fostering professionals who are intrinsically motivated to explore, contest, and further develop meaningful journalistic approaches. This approach is geared towards sustainability rather than novelty and challenges the concept of innovation within data journalism as necessarily groundbreaking: by coherently implementing a plethora of tools and creating a resilient infrastructure for a widespread use of technology, the editor in charge at taz is undoubtedly committed to innovation within the scope of her newsroom. Her goal was to train taz's reporters coherently and create an infrastructure capable of supporting them in developing an entirely new skillset, rather than an individual story.

Transdisciplinarity

Bringing professional figures from domains such as academia, computer science, and design is regarded as key to deal with unusual techniques and tools within complex data journalism projects (Gynnild 2014; Porlezza and Splendore 2019; Appelgren, Linden, and van Dalen 2019). Journalists at taz were not the exception: the possibility of working in a transdisciplinary setting was welcomed enthusiastically from the beginning. Similarly to what has been discussed in previous research regarding data journalists' innovation import in the newsroom – both in German specific contexts (Stalph, Hahn, and Liewehr 2022) and internationally (e.g., de-Lima-Santos 2022): whereas flat cooperation between the parties was warmly welcomed and desired, the journalistic perspective and methodologies were largely favoured while other approaches were subordinated to more practical and mundane tasks. This difference translates directly into disparity of

ownership, meaning that neither the programmer involved in the team nor I had agency in shaping the news-making process. This obstacle to team work contributes again to the interpretation of data journalism epistemology as a form of power constituted by multiple concurring relations (Foucault 1978). When cooperation is at its early stages, external actors are pushed into a subordinate position despite sharing journalists' values and propositions (Baack 2018). Newcomers tend to shield their editorial decisions and operate through traditional methods with little consideration of new potential approaches. Gradually, these different mindsets become part of reporters' will to truth, allowing them to switch between epistemological approaches when reporting with data. This change of position resonates to what has been found by Borges-Rey (2020): journalists actively switch between a traditional and data-driven epistemology when working on data journalism projects. This transition is not seamless and is often a source for frustration rather than creativity enhancing, as previously suggested by Parasie (2015). This finding confirms the risks as well as the importance of conducting transdisciplinary work. On the one hand, technologists and designers struggle in gaining trust and vouch for their positions to be considered as important. On the other hand, by being exposed to new perspectives and methods, journalists broaden the horizons of what shall be known and the material investments needed to produce knowledge through data.

Conclusion

The uniqueness of this specific case allowed me to reflect on the clash between different types of epistemology and their close connection to power as the sum of constraints and boundaries that define social domains. Findings highlight the complexity of introducing data journalism within newsrooms' culture. For newcomers constant negotiation among various aspects is required. First they need to shape their way of approaching data, then they need to be able to accommodate new technological tools, lastly they must learn to collaborate with experts from other fields. Clearly, implementing successful data journalism as a current newsroom practice is not solely dependent on economic resources, but it is also largely linked to how involved practitioners understand its epistemology and leverage its constraints in their favour.

Taz is a cooperative owned company with a peculiar business model and organisational structure. Whereas their high profile journalistic work matches the ones produced in traditional newsrooms, its practices and business culture are at times different, potentially leading to difficult to compare findings. However, the different pace and professional constellations offer an interesting overview of alternative newsroom structures and their study sheds a light on internal and subterranean processes of innovation and adaptation to the current news market. Another possible limitation resides in the deeply qualitative nature of this study. Aside from interviews, I used auto-ethnography extensively and embedded my perspective as a valid source of knowledge, especially when intertwined with the other informants' perspective. From a methodological point of view, they have been conducted following existing literature to avoid any form of self-celebration or pity. The deep connection between human actors and their understanding of their profession and its boundaries makes the employment of such close-reading techniques valid and worthwhile. Lastly, given the timeframe for the project (December 2020 – October 2021) – coinciding with the COVID-19 Pandemic – almost all meetings

and interviews took place online and work has been conducted remotely. This aspect potentially plays an important influence not only on how the project has developed, but also on how the role of data journalists might have been perceived within newsrooms. However, since the focus of this study are newcomers, questions of influence could not be explored and their discussion remain outside of its scope.

Aside from limitations, this study has focused on three particular aspects of data journalism: data, access to technology, and transdisciplinarity. These three are possibly the most straightforward concepts connected to data journalism. They have been discussed at large by many other scholars and they are relatively easy to observe as they get mobilised during team work. Future work should focus on systematically surveying other key aspects of data journalism. This survey should take into consideration both previous academic contributions, to determine how other scholars observed the unfolding of data journalism and its epistemology, as well as observing professional environments where such concepts are mobilised on a daily basis. Ultimately, this research is a step forward in highlighting the nature of data journalism epistemology as a form of power that requires infrastructural and individual changes to be leveraged. Foucault's definitions of power and will to truth have been adopted as a theoretical lens to frame newcomers' struggle in accessing data journalism epistemology. In this sense, a more systematic use of discourse analysis could play a fundamental role in tying these concepts even closer to data journalism epistemology. The reported results focus on the tensions originating from observing the professional struggle of newcomers becoming data journalists. Each tension is discussed both in individual and structural terms to make the transition process between the different epistemologies stand out. The distance between wanting to do data journalism and actually working as data journalists is not easy to bridge. When newcomers are capable of tapping into data journalism epistemology they become able to build their epistemic authority. For reporters it takes courage and dedication to inscribe themselves in the practice. Newsrooms need to undergo fundamental changes to accommodate data journalism and this is not exclusively linked to financial aspects, but also to cultural and professional ones.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

Funding

This project was funded by a grant from the Bundesministerium für Bildung und Forschung (Project: VIDAN – Visuelle und dynamische Arrangements von Nachrichten).

ORCID

Francesca Morini http://orcid.org/0000-0002-5677-8175

References

Ahva, L. 2017. "How Is Participation Practiced by "In-Betweeners" of Journalism?" *Journalism Practice* 11 (2-3): 142–159. DOI: 10.1080/17512786.2016.1209084.

Appelgren, E., C. G. Lindén, and A. van Dalen. 2019. "Data Journalism Research: Studying a Maturing Field Across Journalistic Cultures, Media Markets and Political Environments." *Digital Journalism* 7 (9): 1191–1199. DOI: 10.1080/21670811.2019.1685899.

Appelgren, E., and G. Nygren. 2014. "Data Journalism in Sweden: Introducing new Methods and Genres of Journalism Into "old" Organizations." *Digital Journalism* 2 (3): 394–405. doi:10.1080/21670811.2014.884344.

Appelgren, E., and R. Salaverría. 2019. "Interacting, but not Contributing: Fruitless News Crowdsourcing in Spain." *Estudios Sobre El Mensaje Periodístico* 25 (2): 639–655. doi:10.5209/esmp.64792.

Baack, S. 2018. "Practically Engaged." *Digital Journalism* 6 (6): 673–692. DOI: 10.1080/21670811.2017.1375382.

Borges-Rey, E. 2016. "Unravelling Data Journalism: A Study of Data Journalism Pratice in British Newsrooms." *Journalism Practice 10* (7): 833–843. doi:10.1080/17512786.2016.1159921.

Borges-Rey, E. 2020. "Towards an Epistemology of Data Journalism in the Devolved Nations of the United Kingdom: Changes and Continuities in Materiality, Performativity and Reflexivity." *Journalism* 21 (7): 915–932. doi:10.1177/1464884917693864.

Bounegru, L., and J. Gray, eds. 2021. *The Data Journalism Handbook: Towards a Critical Data Practice*. Amsterdam University Press. doi:10.5117/9789462989511.

Coddington, M. 2015. "Clarifying Journalism's Quantitative Turn." *Digital Journalism 3* (3): 331–348. doi:10.1080/21670811.2014.976400.

de-Lima-Santos, M. F. 2022. "Setting an Agenda to Tackle Environmental Issues with Data and Collaboration." *Journalism Practice* 16 (2-3): 540–560. doi:10.1080/17512786.2022.2031256.

de-Lima-Santos, M. F., and L. Mesquita. 2021. "The Strategic Value of Data Journalism." In *Journalism, Data and Technology in Latin America*, edited by R. Salaverría and M. F. de-Lima-Santos, 97–136. Cham.: Palgrave Macmillan.

Dowling, D., and T. Vogan. 2015. "Can We "Snowfall" This?" *Digital Journalism* 3 (2): 209–224. doi:10.1080/21670811.2014.930250.

Ekström, M., A. Ramsälv, and O. Westlund. 2022. "Data-driven News Work Culture: Reconciling Tensions in Epistemic Values and Practices of News Journalism." *Journalism* 23 (4): 755–772. doi:10.1177/14614448221150439.

Eriksson, P., E. Henttonen, and S. Meriläinen. 2012. "Ethnographic Field Notes and Reflexivity." In *An Ethnography of Global Landscapes and Corridors*, edited by Loshini Naidoo, 9–22. IntechOpen.

Fink, K., and C. W. Anderson. 2015. "Data Journalism in the United States: Beyond the "Usual Suspects"." *Journalism Studies* 16 (4): 467–481. doi:10.1080/1461670X.2014.939852.

Foucault, M. 1971. "Orders of Discourse." *Social Science Information* 10 (2): 7–30. doi:10.1177/053901847101000201.

Foucault, M. 1975. *Discipline and Punish. The Birth of the Prison*.

Foucault, M. 1978. *The History of Sexuality*. New York: Pantheon Books.

Gershon, N., S. G. Eick, and S. Card. 1998. "Information Visualization." *Interactions* 5 (2): 9–15. doi:10.1109/MCG.1997.595265.

Groll, T. 2012. die tageszeitung (taz). In: mediadb.eu https://www.mediadb.eu/forum/zeitungsportraets/die-tageszeitung-taz.html.

Grubenmann, S. 2016. "Action Research." *Digital Journalism* 4 (1): 160–176. DOI: 10.1080/21670811.2015.1093274.

Gynnild, A. 2014. "Journalism Innovation Leads to Innovation Journalism: The Impact of Computational Exploration on Changing Mindsets." *Journalism* 15 (6): 713–730. doi:10.1177/1464884913486393.

Haim, M. 2022. "The German Data Journalist in 2021." *Journalism Practice* 0: 1–20. doi:10.1080/17512786.2022.2098523.

Hearn, G., J. Tacchi, M. Foth, and J. Lennie. 2009. *Action Research and New Media*. Hampton Press.

Hepp, A., and W. Loosen. 2021. "Pioneer Journalism: Conceptualizing the Role of Pioneer Journalists and Pioneer Communities in the Organizational Re-figuration of Journalism." *Journalism* 22 (3): 577–595. doi:10.1177/1464884919829277.

Heravi, B. R., and M. Lorenz. 2020. "Data Journalism Practices Globally: Skills, Education, Opportunities, and Values." *Journalism and Media* 1 (1): 26–40. doi:10.3390/journalmedia1010003.
Kosterich, A., and M. S. Weber. 2018. "Transformation of a Modern Newsroom Workforce." *Journalism Practice* 4: 431–457. doi:10.1080/17512786.2018.1497454.
Lewis, S. C., and N. Usher. 2014. "Code, Collaboration, and the Future of Journalism: A Case Study of the Hacks/Hackers Global Network." *Digital Journalism* 2 (3): 383–393. doi:10.1080/21670811.2014.895504.
Lewis, S. C., and O. Westlund. 2014. "Big Data and Journalism: Epistemology, Expertise, Economics, and Ethics." *Digital Journalism* 3 (3): 447–466. doi:10.1080/21670811.2014.976418.
Lincoln, Y. S., and E. G. Guba. 1985. *Naturalistic Inquiry*. London: Sage Publications, Inc.
Marshall, J. 2001. "Self-reflective Inquiry Practices." In *Handbook of Action Research: Participative Inquiry and Practice*, edited by P. Reason, and H. Bradbury, 335–342. London: Sage Publications Inc.
McIntosh, C.. 2022. *Data in Cambridge Essential Dictionary*. Cambridge University Press. https://dictionary.cambridge.org/dictionary/essential-british-english/data.
Mutsvairo, B. 2019. "Challenges Facing Development of Data Journalism in non-Western Societies." *Digital Journalism* 7 (9): 1289–1294. doi:10.1080/21670811.2019.1691927.
Naidoo, L. 2012. An Ethnography of Global Landscapes and Corridors https://openresearchlibrary.org/content/903563ac-c1ad-4835-bbbe-ce60a6ee89a0.
Parasie, S. 2015. "Data-Driven Revelation?: Epistemological Tensions in Investigative Journalism in the Age of "big Data"." *Digital Journalism* 3 (3): 364–380. doi:10.1080/21670811.2014.976408.
Parasie, S., and E. Dagiral. 2013. "Data-driven Journalism and the Public Good: "Computer-Assisted-Reporters" and "Programmer-Journalists" in Chicago." *New Media & Society* 15 (6): 853–871. doi:10.1177/1461444812463345.
Picha Edwardsson, M. 2022. "Automating the Search and Selection of News in Municipal Registries: Experiences from a Swedish Participatory Action Research Project." *Journal of Applied Journalism & Media Studies* 00(00): 1–21. doi:10.1386/ajms_00082_1.
Porlezza, C., and S. Splendore. 2019. "From Open Journalism to Closed Data: Data Journalism in Italy." *Digital Journalism* 7 (9): 1230–1252. DOI: 10.1080/21670811.2019.1657778.
Segel, E., and J. Heer. 2010. "Narrative Visualization: Telling Stories with Data." *IEEE Transactions on Visualization and Computer Graphics* 16 (6): 1139–1148. doi:10.1109/TVCG.2010.179.
Siapera, E., and L. Papadopoulou. 2016. "Entrepreneurialism or Cooperativism?" *Journalism Practice* 10 (2): 178–195. DOI: 10.1080/17512786.2015.1125760.
Stalph, F., and E. Borges-Rey. 2018. *Data Journalism Sustainability, Digital Journalism* 6 (8): 1078–1089. DOI: 10.1080/21670811.2018.1503060.
Stalph, F., O. Hahn, and D. Liewehr. 2022. "Local Data Journalism in Germany: Data-Driven Reporting Amidst Local Communities and Authorities." *Journalism Practice* 0: 1–20. doi:10.1080/17512786.2021.2019089.
Stringer, E. T. 2013. *Action Research*4th ed. London: SAGE Publications, Inc.
Szilárdi, R., and M. Tőke. 2019. "Crowded Intersections: Transdisciplinary Tendencies in Social Sciences and the Humanities." *Belvedere Meridionale* 31 (4): 53–64. doi:10.14232/belv.2019.4.5.
Usher, N. 2020. "News Cartography and Epistemic Authority in the Era of Big Data: Journalists as Map-makers, Map-users, and Map-subjects." *New Media & Society* 22 (2): 247–263. doi:10.1177/1461444819856909.
Wagemans, A., and T. Witschge. 2019. "Examining Innovation as Process: Action Research in Journalism Studies." *Convergence* 25 (2): 209–224. doi:10.1177/1354856519834880.
Yin, R. K. 2004. *The Case Study Anthology*. SAGE Publications, Inc.
Zamith, R. 2019. "Transparency, Interactivity, Diversity, and Information Provenance in Everyday Data Journalism." *Digital Journalism* 7 (4): 470–489. doi:10.1080/21670811.2018.1554409.