Towards an epistemology of data journalism in the devolved nations of the UK: Changes and continuities in materiality, performativity and reflexivity

Eddy Borges Rey, University of Stirling

Abstract

This article outlines a general epistemological framework of data journalism in the devolved nations of the UK. By using an original model based on three conceptual lenses—materiality, performativity and reflexivity—this study examines the development of this form of journalism, the challenges it faces, and its particularities in the context of Scotland, Wales and Northern Ireland. This research therefore offers unique insights from semi-structured interviews with data journalists and data editors based at, or working as freelancers for, the mainstream news organisations of these regions. The results suggest that data journalism in these devolved nations displays a distinctive character just as much as it reinforces the norms and rituals of the legacy organisations that pioneered this practice. Whilst various models of data exploitation are tested, regional data journalists creatively circumvent generalised organisational struggles to lay the groundwork for their trade and professional community.

Keywords

Data journalism, hyperlocality, FOIA, materiality, performativity, reflexivity, devolved nations of the UK, epistemology

Introduction

Since the term *big data* began to gain traction in 2010-11, a growing number of analysts, scholars and professional journalists have endeavoured to both make sense of its complex nature and understand its impact on contemporary journalism practice. Prior to the current scholarly excitement surrounding *data journalism*, Hamilton and Turner (2009: 2) defined computational journalism as a 'combination of algorithms, data, and knowledge from the social sciences to supplement the accountability function of journalism'. Five years later, when the foundations of the study of data journalism practice have been

laid down by a series of national and organisational case studies (Parasie and Dagiral, 2013; Fink and Anderson, 2015; Parasie, 2015; Karlsen and Stavelin, 2014; Appelgren and Nygren, 2014; De Maeyer et al., 2015; Tabary et al., 2015; Borges-Rey, 2016), Lewis and Westlund observe 'For journalism, big data embodies emerging ideas about, activities for, and norms connected with data sets, algorithms, computational methods, and related processes and perspectives tied to quantification as a key paradigm of information work' (2015: 449). Despite the apparent similarity of both conceptualisations, scholars and practitioners alike have realised that an all-encompassing working definition of data journalism is rather difficult to achieve[1], as it would require normalising tensions between the competing ontological features displayed by a variety of journalistic methodologies that deal with data. As the epistemology of journalism is pervaded by data cultures, it moves back and forth between reporting and programming, storytelling and coding, hypothesis-driven inquiry and data-driven inquiry, journalistic thinking and computational thinking, interactivity and staticity, visuality and textuality, centrality and locality, individuality and collaboration, or human-driven paradigms and machine-driven paradigms, to name a few. Perhaps it is precisely the ever-changing polarity between these tensions that has prompted the proliferation of a series of overlapping terms, definitions and ways of understanding *data journalism* that I will use later on in this article to construct my epistemological framework.

Amid the current academic emphasis on national analyses, this article offers, for the first time in the field, a hyperlocal study that delineates a general epistemological framework for data journalism in the devolved nations of the United Kingdom. By using notions of materiality, performativity and reflexivity as conceptual lenses, the article advances the thesis that Scotland, Wales and Northern Ireland present a set of distinctive features that at times make data journalism practice unique and therefore deserving of its own epistemological framework. However, this emerging epistemology will also have to recognise that there are as many changes as continuities in the way data journalism practice is mediated in those nations.

To develop this argument, the research presents a novel analysis which is articulated around two predominant approaches constructed on the basis of previous research on the epistemologies of data journalism (Parasie and Dagiral, 2013; Gynnild, 2014; Coddington, 2015; Parasie, 2015). On this basis, I refer to the first approach as *'newshound',* or the traditional journalistic ways of handling and

engaging with data; and the second one as *'techie',* which designates an emergent journalistic approach to data based on more computational logics and mindsets.

News media in the devolved nations of the UK: locality, struggle and identity

Undoubtedly, devolution[2] has played a key role in shaping the distinctiveness of the news media in Scotland, Wales and Northern Ireland (Schlesinger, Miller and Dinan, 2001; McNair, 2006; Cushion, Lewis and Groves, 2009). However, prior to devolution, each nation's indigenous media displayed distinctive epistemological features that were moulded by a series of significant historical developments. Jones (1993), for instance, argues that in the 19th century, the Welsh press was compartmentalised by advertisers into economic regions that grew unevenly, developing not only different types of societies but also different forms of journalism. Similarly, Rolston remarks that the rules and rituals of indigenous media in Northern Ireland were shaped by two decades of nationalist 'troubles' (1991:2). Scotland has also displayed a distinctive media identity that was influenced by intellectual movements such as the Enlightenment, which created "a public sphere of exceptional richness" (McNair, 2006) and an audience avid for local content (McNair et all, 2010).

But throughout their history, perhaps the most significant incentive for these 'national' news outlets to thrive was the recurrent marginalisation of Scottish, Welsh and Northern Irish news by London-based news organisations (Jones, 1993; Cushion, Lewis and Groves, 2009; Conboy, 2011). This attempt to homogenise a British news agenda through the prism of the London-based media, together with other factors, led indigenous regional media to foster and reinforce a unique identity.

Nowadays, the three devolved nations have in common an indigenous press that fervently serves a local news agenda to cope with international and London-based competition; whilst centralised broadcast corporations tend to approach "alternative national audiences within what was still presented as an unproblematically homogenous nation" (Conboy, 2011: 187). Although still in its infancy, local television has shifted towards a more regional news agenda.

In this media context, with collective, regulated, privileged access to local government (Schlesinger, Miller and Dinan, 2001), regional newspapers act as a counteroffensive against sustained declines in circulation (Dekavalla, 2015), global economic competition, technologically driven change, and supranational processes of political and economic integration (Schlesinger, 1998: 72).

An epistemological cartography of data-driven journalism

During its relatively brief academic history, data-driven journalism has been commonly described through a range of competing and overlapping conceptualisations including computer-assisted reporting (CAR), data journalism, precision journalism, computational journalism, programming journalism, and algorithmic journalism.

As a consequence, attempts to articulate a general ontology of data journalism have been unsuccessful. Indeed, one of the long-debated topics in the field has been what *data-driven journalism* actually entails. Is it a reinvigorated form of investigative journalism and its novelties mere additions to the traditional workflow of journalists? Or is it, on the other hand, an innovative combination of the methods, philosophies and logics of computer science and journalism that explores the use of pioneering technologies in redefining the procedures of gathering, producing, presenting and disseminating news?

As a means to clarify this epistemological ambiguity, in this section I will offer an overview of the following four seminal explorations of the epistemology of data journalism: Parasie and Dagiral (2013), Gynnild (2014), Coddington (2015), and Parasie (2015), to identify their commonalities. Encompassing a wide range of perspectives, these studies largely distinguish between two approaches. The first is subordinate to axiomatic professional norms and sees data as one of the many evidentiary components of news. The other is able to switch between journalistic and computational mindsets to leverage the best from both worlds in a quest to create ground-breaking editorial products. Hence, in summarising the results of these studies, I situate their epistemological emphases on a spectrum that features an opposing approach at either end, and that I have called the *newshound* approach and the *techie* approach (see figure 1).

Beginning with Parasie and Dagiral (2013), I place their first epistemological proposition near the *newshound* approach, which they see as a continuation of the muckraking journalistic tradition whereby reporters add value to data when they find a hidden story within it (2013: 859). Later, in his second piece, Parasie refines this traditional model by diversifying it into the epistemologies of investigative journalism and CAR (2015: 367-68). Here, I also place Gynnild's newsroom approach, which replicates the normative methods employed by journalists to address quantitative data (2014). Within the framework of that approach, Gynnild includes the modalities of CAR and precision

journalism, whilst she places data journalism somewhere between her newsroom approach and what she calls the entrepreneurial approach (2014: 717-19). In addition, I place Coddington's modality of CAR near the *newshound* approach. Utilising a prism consisting of four professional and epistemological dimensions defined by contrasting tendencies—professional expertise versus networked information, transparency versus opacity, targeted sampling versus big data, and active versus passive visions of the public—Coddington views CAR as favouring normative paradigms, such as professional expertise, opacity, target sampling and a passive vision of the public (2015).

Towards the other end of this spectrum, I place Parasie and Dagiral's second epistemological proposition closer to the *techie* approach. Rooted in computing logics, this model is fostered by a progressive collaboration amongst journalists, programmer-journalists and civil actors with computing proficiency who seek to strengthen government accountability and citizen participation by releasing public data (2013: 860-61). In his newest article, Parasie further develops this epistemology. As investigative and computer-assisted reporters collaborate with their colleagues who undertake a *data-driven approach*, tensions arise when both groups produce beliefs justified by the premises of either a hypothesis- or a data-driven path, demonstrating how different the approaches of journalists and data analysts can be when producing data-driven news stories (2015).
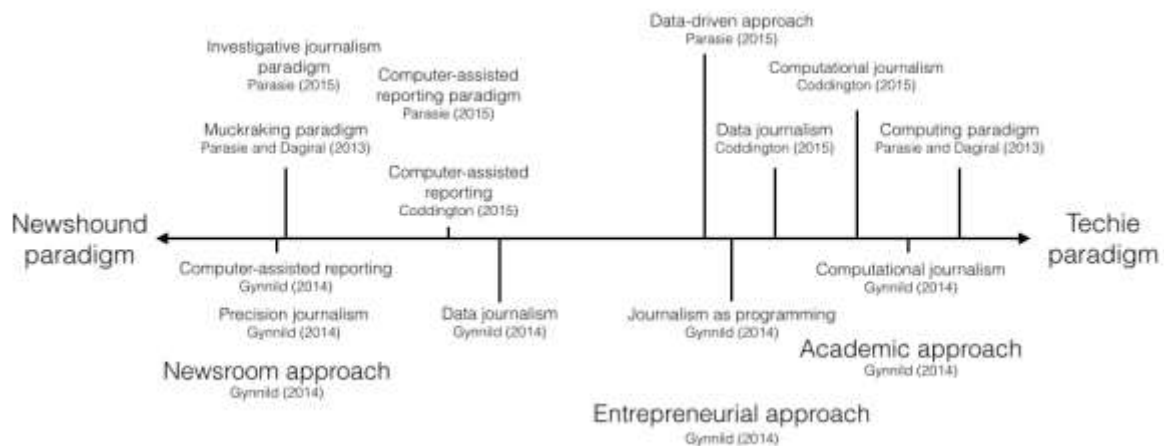

Fig. 1 – The epistemologies of data-driven journalism practice

I also place Gynnild's entrepreneurial approach closer to this end of the spectrum. Gynnild notes that the most significant innovations in journalism are driven by forms of *journalism as programming* outside the rigid structures of news organisations, placing databases at the epicentre of news

attention (2014: 720-21). I place her academic approach even closer to the *techie* end, whereby multidisciplinary researchers contribute to developing modalities of computational journalism strongly mediated by computational thinking[3] (2013: 722). Although Coddington observes that the three modalities he analysed share similar epistemological roots and are therefore not mutually exclusive (2015: 333), I also place his modalities of data and computational journalism towards the *techie* end, as they reveal behaviour that contrasts with CAR, in addition to favouring, to somewhat disparate degrees, networked information, transparency, big data and an active conception of the public.

As each of these epistemological propositions suggests, the primary definers of data journalism are mediated by a constant interplay between an approach that is deeply rooted in journalistic conventions and an approach that is increasingly reliant on computational processes and logics. For the purposes of the forthcoming material, performative and reflexive analyses, I consider the cases studied within the context of either a *newshound* or *techie* approach.

Method

To find out whether there is a distinctive epistemology of data journalism in the devolved nations of the UK, this article explores the current state of affairs of data journalism practice in Scotland, Wales and Northern Ireland to determine a) what the particular idiosyncrasies of data journalism in these devolved nations are, b) how knowledge claims are produced and disseminated by data journalists, and c) the challenges faced by the trade when legitimising their knowledge claims. By means of a novel material, performative and reflexive examination, the article investigates how data mediates the outputs, tools, techniques, strategies and mindsets of data journalists and data editors working for news organisations in these devolved nations.

This qualitative research draws on a combination of nine semi-structured interviews with key informants based at or working as freelancers for STV, BBC Scotland, the Scotsman, the Glasgow Herald, thedetail, the News Letter, and Trinity Mirror (Regionals). For the overall analytical framework, this research uses the notions of materiality, performativity and reflexivity as conceptual lenses.

The number of news organisations in the devolved regions with the infrastructure required to employ data journalists has been so far relatively small, which made the identification of key informants relatively straightforward. Informants were identified by either a localised internet search of data-driven stories, data journalists, or data units in these nations, or by asking the informants that were

already involved about other data journalists working in the area. After the identification of key informants, most of the population of professional data journalists working for mainstream news organisations in Scotland, Wales and Northern Ireland, were interviewed. The group of key informants interviewed in this research includes seven data journalists or journalists working with data (coded as DJ1- DJ7) and two data editors (coded as DE1 and DE2). The geographic segmentation of the informants is as follows: four informants were based in Scotland (all DJs), two informants were based in Northern Ireland (a DJ, and a DE) and two informants were based in Wales (two DJs). There is an additional divergent case that was incorporated into the group of informants for specific reasons; namely the data editor of Trinity Mirror who coordinates the regional data journalism operations from Manchester. The primary function of his unit is to generate data-driven resources and news stories for all of Trinity Mirror's English titles as well as for the *Western Mail*, *South Wales Echo*, and *Wales Online* in Wales; the *Daily Mirror* in Belfast and the *Irish Mirror* in Dublin; and (occasionally) the *Daily Record* and affiliated local titles in Scotland[4]. At the time the interviews were conducted, one of the data journalists who provided information on the Welsh context was based in Cardiff, working for her data journalism unit remotely; whilst the other had been relocated a week before the interview to Trinity Mirror's London offices.

Secondary research was undertaken to construct the analytical frameworks of the material (See Coole and Frost, 2010; Miller, 2005; Van der Tuin and Dolphijn, 2012), performative (See Broersma, 2010) and reflexive (See Guillaume, 2002) lenses. Next, recorded semi-structured interviews lasting 83 minutes on average were conducted face-to-face or via Skype between January 2014 and August 2015. Interviews focused on different aspects of the data news production process, i.e., techniques and strategies used by data journalists to filter databases and datasets, to identify and isolate newsworthy elements from numerical data, to translate mathematical abstractions into components of a news story, and to present data in an engaging way to the audience. The empirical data gathered was analysed in the context of a study on data journalism in the UK (Borges-Rey, 2016) and a series of national and institutional case studies in the U.S. (Parasie and Dagiral, 2013; Fink and Anderson, 2015; Parasie, 2015), Norway (Karlsen and Stavelin, 2014), Sweden (Appelgren and Nygren, 2014), Belgium (De Maeyer et al., 2015) and Canada (Tabary et al., 2015) to offer wider comparative insight. The resulting data was then analysed from the point of view of the following three thematic domains:

1) the materials of data journalism, 2) the performative routines and values of data journalism, and 3) the reflexivity of data journalism.

Results

Data journalism in the devolved nations of the UK displays a distinctive character just as much as it reinforces the norms and rituals of the legacy organisations that pioneered its adoption in the UK and abroad during the last decade. After contrasting the data gathered through my interviews against the existing research on the field, I provide a theorisation based on an interpretation of the testimonies of my informants. During the interviews, it became evident that data journalists and editors, although fluent in data-related practices, were still trying to make sense of this emerging form of journalism in the context of their particular news organisations. With varying levels of complexity, each informant approached the practice to the best of their abilities, with a very critical understanding of what data journalism meant to them, the benefits and pitfalls of the trade, and how to make it work in their regions.

In this sense, the following sections will show very different media organisations, professionals and ways of engaging with the trade. For instance, legacy organisations such as *BBC Scotland* or *STV* perform very individualistic data journalism routines –with only one or two data journalists– whereas *Trinity Mirror* has a data journalism team functioning as a centralised datafied news wire that produces generic content for regional reporters. Non-legacy actors play also a key role in shaping 'national' data journalism. For example, freelancers in Scotland provide legacy organisations like *The Scotsman* or *The Glasgow Herald* the opportunity to publish pieces of data journalism that they cannot afford in-house.

The first conceptual lens: Data journalism materiality

In this section I argue that data materiality serves as the breeding ground of data journalism's performativity and reflexivity. To corroborate this argument, I aimed to see whether informants understood data beyond its common reductionist characterisation as one of the many ingredients of the recipe we call news, and hence controllable, inactive and subordinate to human agency.

The materiality of FOI-driven data

If the material identity of any object can be understood in terms of the historical process that originated it (see DeLanda's interview in Van der Tuin and Dolphijn, 2012: 39), the testimonies I gathered replicated extant studies of the field, which report that data commonly originates through an in-house collation managed by news organisations, or is acquired through subscription, Freedom of Information (FOI) requests, open data regimes, and/or leaks.

Evidence from my interviews suggests that data's most significant historic definer was the Freedom of Information Act (FOIA). The vast majority of informants agreed that most of the data they use originated within the public sector and was obtained by means of FOI requests. The FOIA scheme thus largely infuses the material character of the datasets used as evidence by data journalists in Scotland, Northern Ireland and Wales. In this respect data is understood by most of the informants to be generated within the public domain (DJ1, DJ2, DJ3, DJ4, DJ5, DJ6, DJ7, DE1, DE2), and in some cases influenced by the agendas of politicians, press officers, and other spin doctors (DJ1, DJ2, DJ6). Data is also perceived to be restricted through the bureaucratic mechanisms of public bodies (DJ2), which in some cases, after prolonged negotiation, provide the data in a non-machine-readable format (DJ1, DE1) making it difficult to use.

Other sources of data were also discussed by informants, such as tip offs by statisticians or record officers working for public authorities (DJ1, DJ6); scraping websites (DJ3, DJ7, DE2); third-sector organisations or private companies willing to grant access to their databases (DJ7, DE1); leaks (DJ5); and data feed subscriptions or data licenses.

Data beyond matter

Bearing in mind Miller's (2005: 3) distinction between things as artefacts, and things beyond the dualism of subjects and objects, the evidence suggested that data was predominantly assumed by informants to be a material artefact that functions as either an evidential input for stories or as data-driven output. Perceived as an inanimate good, data is claimed by data journalists once they have found it within the convoluted entanglement of datasets held by public bodies, and subsequently placed in the formulaic structure of the news item (DJ1):

> 'I think I claimed that story and I'll do it every year now, and it gets followed up by other local
>
> press, that's the kind of way things happen. [DJ4] gets all the environmental stories. There's no

point really in trying to get those, he picks on the freedom of information request, works well

with the data, and sources send him stories as well, so he's really claimed that patch' (DJ1)

In the context of the devolved nations, then, the material identity of data appears to be heavily

dependent on the human agency of either journalists, or experts who make data understandable after

contextualising it (DJ1). Now, following my initial argument that data may be understood beyond the

notion of data artefacts, signs of materiality beyond objectifying views were identified less frequently

across the testimonies of the informants. As journalists engaged with data daily, there was a general

unawareness about the extent to which data can mediate work routines. Regardless, informants

acknowledged that data enabled them to generate exclusives (DJ5, DJ6, DJ7, DE2) in a process

whereby data journalists, whilst recognising data's capacity to make their stories more accurate and

rigorous, simultaneously were capable of challenging the data by looking for flaws within datasets

(DE1, DJ2). Editors were also able to shape data into editorial products capable of either being

repurposed or reused to recoup their initial investment, or enabling income generation and audience

engagement. A couple of interviewees implied that the most active embodiment of data was perhaps

the web-metrics that inform editorial decision-making (DE2).

Datafied agency

Journalistic algorithms, hybrid entities that, as Anderson puts it, are 'neither entirely material, nor are

they entirely human' (2013: 1016), are becoming increasingly popular in academic research.

Nonetheless, automated systems, shaped as custom-made entities that only elite news organisations

can afford, are generally infrequent in the context of the devolved nations. In agreement with previous

research, informants remarked that smaller, local news organisations must normally rely on generic

third-party solutions, which, luckily, are freely accessible through open-source regimes but are not

always compatible with the existing architecture of their newsrooms, thus hindering the aptness of the

algorithmic organism for the ecosystem in which it performs. All the informants declared that they

used at least one form of software automation during their working practices. However, only two data

journalists and one data editor working for legacy organisations expressed their familiarity with web

scraping (DJ3, DJ7, DE2), albeit at different levels of expertise. Two of them use the technique on a

regular basis during their news reporting (DJ3, DJ7). These are custom-made scripts written by the

journalists to solve very contextualised problems; therefore, they tend to be effective and helpful in optimising their workflow.

The remainder of the informants displayed a level of inexperience with coding that made them dependent on collaborations with programmers or a few generic software solutions that were limited by the universality of their design. Notably, these informants displayed a remarkable ability to circumvent the constraints of generic software by seeking additional applications that filled the original gap. Spreadsheet applications were generally used for instrumental data analysis (DJ1, DJ5, DJ6). Visualisation tools were used for either presenting results in visual format (DJ1, DJ3) or as visual aids to identify trends during the analysis stage (DJ5). Some informants used online apps to programme automated tasks and generate potential leads (DJ1) or to clean messy data (DJ3).

To summarise, as journalists in the devolved nations of the UK engage with data in a series of power negotiations where both journalists and data switch back and forth between their roles as enabler or enabled, my analysis suggests that informants seemed to subscribe more firmly to a *newshound* approach, whereby they would normally think of data as an element that is subordinated to the news reporting process and acquires significance within the context of a story. Conversely, evidence of a *techie* approach was less frequent.

The interviews suggest that data in these nations is instrumentally handled and heavily negotiated by the politics of public organisations and the human agency of data journalists, which, at this early stage, arguably points to the initial prevalence of the *newshound* approach. The next section will explore in detail the ways in which data materiality intermediates the meanings data journalists attach to their working practices.

The second conceptual lens: Data journalism performativity

Over the past few years, research has shown the large extent to which data journalism is progressively legitimising its own professional norms and discourses. In this section, therefore, I will examine those norms and discourses in order to understand how they facilitate data journalism performativity.

In accordance with previous research[5], the analysis of the interviews suggested that in order to persuade their audiences of the veracity of their accounts (Broersma, 2010: 17-18), journalistic

authority endured as the imperative requirement to perform within the trade notwithstanding the apparent importance of data and computation. Informants unanimously stated that the most significant professional value required to legitimise their knowledge claims was traditional journalistic skills. This was to be expected, as all of the informants received formal training as journalists in one way or another, and tended to adhere to the norms and routines of legacy organisations –despite functioning as online-only platforms or freelancers. Data journalism in the devolved nations of the UK thus remains heavily driven by journalists, if compared with London-based organisations, where, non-legacy actors such as programmers, statisticians, or graphic designers have a more active role in the production of news (Borges-Rey, 2016). There were however a number of cases where data journalists had previous professional expertise that complemented their current professional profiles. One of the informants, for instance, has a bachelor's degree in History and a Masters in Computing Science, and after working as developer for 10 years he completed a Masters in Journalism. Similarly, another informant remarked 'I come from a community engagement sector' (DJ2), which evidences the hybrid character of the professional background of some of the data journalists interviewed.

Journalistic abilities mirrored those recorded in findings from previous research, and were deemed by informants as particularly valuable for a) finding the right data for the story or the story behind the data; b) for negotiating the acquisition of data denied by public officials or provided in a non-machine-readable format (DJ1, DJ2, DE1); c) for scrutinising the accuracy of the data gathered and removing potential spins (DJ1, DJ6); d) for finding the human interest in—and provide context to—data through human testimony (DJ1, DJ2, DJ4, DJ6); e) for translating data by filtering, selecting and identifying the most relevant trends therein to attract readers (DJ1, DJ2, DE1); and f) to engage audiences with innovative forms of storytelling (DJ5). Editorial judgment and *nose for news* remain essential professional values that, on many occasions, override the dictates of data metrics.

Although important, informants indicated that other conventions occupy a secondary place, i.e., that they do not require proficiency in a strict sense. At least one-third of the data journalists interviewed acknowledged having a background in a mathematics-related discipline, further complementing the hybrid character of the trade in the devolved nations. The remainder showed competence in basic statistical tools (DJ1, DJ2, DJ3, DE1). Informants were aware of recurrent statistical errors in news reporting and how to avoid them (DJ3, DJ4); they were also extremely cautious with respect to their

calculations (DE1), corroborating with public bodies what types of formulas and algorithms were used to process data (DJ3).

Rigour in their methods also reinforced the perceived soundness of their reports. During the interviews, an informant stressed that, in order to remain trustworthy in the public eye, scientific rigour was imperative to regain the trust that journalists lost after the Leveson Inquiry: '[W]e as journalists need to show our work, just as mathematicians and scientists, and show sources, and I think that might help to raise trust in journalism as a profession' (DJ1). Another data journalist highlighted the difficulty in harmonising two very different types of methods and narratives:

> 'So you have to always find this compromise between scientific accuracy and something you are trying to get across. So I just tell the story because it illustrates the sort of negotiation and process that has to go on, taking a sort of raw academic, scientific paper and translating it into a table that would mean something in a newspaper' (DJ4).

Informants also stressed how important it was to remain in contact with the local sources that generate or interpret raw data. In particular cases, this relationship served to verify the soundness of data journalists' methods and calculations throughout the process, identifying and removing errors or misinterpretations before submitting the piece for editorial perusal (DJ3, DJ4, DE1). Although the role of non-legacy actors, such as local council statisticians, is key to data journalism in Scotland and Northern Ireland –and reporters treasure them as valuable aid– their actual involvement remains both sporadic and peripheral to the core of news production.

It was common practice amongst the interviewed data journalists to explain their use of methods and computing tools as part of their stories. In this regard, the apparent neutrality of computation is a value frequently used to strengthen the perceived accuracy and timeliness of their accounts. Using machines for data collection or synthesis arguably reduces human bias, facilitating the emergence of stories and trends that otherwise might have taken more time to identify (DJ1, DJ5) and, in some cases, can speed up the turnaround time to fit the news cycle (DJ3).

Computation has also assisted the rise of crowdsourced cooperation—the first of a set of values that data journalists use to both generate and explain phenomena simultaneously (Broersma, 2010: 17-18). When asked about the subject, informants mentioned they had attempted crowdsourcing at least once, by identifying local communities of online users that actively aided them in disseminating,

analysing or generating information. However, in most of the cases they seemed to fail to attract numbers significant enough to deem the practice successful. Informants attributed responsibility for this lack of engagement with crowdsourcing to the common denominator of regional audiences, who normally prefer to consume news from newspapers, radio or TV, and tend to be generally disengaged with online news.

In line with previous research[6], open source philosophies behind the notion of crowdsourcing also encouraged informants to pursue extra- or intra-newsroom collaborations. From their testimonies, it became clear that the absence of certain advanced computational skills and/or the restricted access to certain information has compelled informants to embrace these open source ideals and seek internal or external collaboration in their efforts to overcome these limitations. Internally, informants generally reached out to fellow reporters on other beats to seek advice or assistance (DJ6). Moreover, some informants have an essential role in enhancing data literacy within their newsrooms through training programmes (DJ1, DE2). Better equipped units have expanded to incorporate graphic designers or coders as part of the team (DE1, DE2). Informants also stressed that data journalism should ideally be mediated by a constant collaboration between data journalists, graphic designers, programmers and statisticians (DJ1, DJ2, DE1, DE2). Nevertheless, most 'national' legacy organisations could not afford this type of interaction on a regular basis.

Externally, relations with academics, public archivists, records officers and/or statisticians were considered important. Informants also mentioned collaborations with civic organisations and open source initiatives (such as Hacks/Hackers or scraperwiki) that aim to establish partnerships with developers or seek advice from communities of programmers (DJ1, DJ3). Collaborations with libraries, trusts or foundations (DE2) were deemed fruitful when the datasets were collated, curated and maintained by these organisations in a pristine manner. Occasionally, some projects engaged in a partnership with third-sector organisations (DE1) or private companies (DE2, DJ7) that contributed to the project by opening up their databases for journalistic scrutiny. Data-driven stories produced by some of the informants were occasionally covered by local newspapers, thus contributing to the development of the local news ecosystem (DJ1, DJ5). As previously stated, most of these non-legacy actors are involved in delineating the contours of the epistemology of data journalism in the devolved nations of the UK to a degree, as these cases are rather infrequent remaining tangential to the epistemological core of the trade.

One of the most invaluable sources of information and support for these regional data journalists though, is their data journalism community. As the regional contingent remains rather small and data journalists seem to be anchored to their local data allotments, their relationship transcends the boundaries set by brands, and competition is set aside to reinforce ideals involving supportive openness (DE1).

Finally, by default data journalists in these devolved nations strengthen their resilient sense of hyperlocal connection by making use of local data to address issues within their community. As local government engages with the current datafication of social life and opens up its databases for public scrutiny, the constant intermediation between the ideals of transparency and openness and values such as methodological rigour amplifies data journalists' social responsibility towards their communities. This hyperlocal empathy is shown across the board by informants as they passionately address issues including the lack of scrutiny in council meetings and reports (DJ1), governmental data scarcity (DJ1), power abuse (DJ2, DJ6), administrative bureaucracy (DJ2), communal deprivation (DJ1, DE1) or civic disengagement (DJ5, DE2). Whilst access to granular local data generally enhanced the sense of hyperlocal empathy of data journalists in Scotland and Northern Ireland, hyperlocality was achieved in a different way by Trinity Mirror. With the data unit functioning as a centralised wired system, data journalists based in Wales have to cover around two general news topics per day, and produce five to eight stories for each region from these (DJ7). Within this workflow, and despite the potential for devolved data to create strongly local data pieces, it is in many cases the local reporter who makes the general data piece hyperlocal.

In tune with previous case studies[7], the results from the interviews evidenced that data journalists' performativity in the devolved nations of the UK tends to be strongly challenged by internal organisational and editorial pressures, and by the scarce availability of human and material resources, if compared with London-based organisations (Borges-Rey, 2016). On one hand, these issues, which tend to prevail more in small than in large news organisations, have a detrimental effect on the performativity of data journalists who subscribe to the *newshound* approach because they tend to be reliant on collaboration with extra-newsroom agents. On the other hand, the performativity of data journalists who subscribe to a *techie* approach tends to be relatively independent of these issues, and professionals with such specialised performativity tend to be hired by large news organisations for data-related tasks. In this regard, their performativity derives from more advanced

computational undertakings as long as these are not limited by a lack of technological tools and infrastructure.

In terms of the overall professional development of the trade in these devolved nations, fully functioning data units and dedicated data journalism coverage by *the detail* and Trinity Mirror enabled a discreet but healthy data journalism news stream in Northern Ireland and Wales. Notably, despite being the only region with a devolved parliamentary policy for data management, with only one exception, Scotland displayed a degree of editorial hesitancy towards data journalism, which was interpreted by most of the Scottish informants as a barrier to the definitive consolidation of data journalism in Scotland. Staffing and resource limitations were highlighted by informants, which is consistent with extant research on the field. Only one of the data journalists working in Scotland has been officially appointed as a senior data journalist, whereas the others work as freelancers, or generate data stories in their spare time because data journalism is not part of their core remit. Paradoxically, with a public sphere avid for local content and saturated with national and regional titles, it is the Scottish press that fails to recognise the ability of data journalism to generate exclusives from local data that could provide a competitive advantage, better serve a national news agenda, and reinforce a distinctive local identity. The same is true for legacy organisations in Northern Ireland and Wales, which could tap into the wealth of data generated by local authorities and uncover information with extraordinary appeal to local audiences. Unfortunately, with 'national' legacy organisations unable to fully embrace data journalism as a normative practice in their newsroom, and skills such as numeracy, data literacy and basic notions of coding becoming essential for the trade, smaller local news organisations remain incapable of affording a practice that could have remarkable hyperlocal potential.

Third conceptual lens: Data journalism reflexivity

Thus far, I have explored the ways in which data materiality—both as an artefact and beyond its boundaries—mediates data journalism performativity. In this section I examine the subsequent process by which the interplay between data materiality and data journalism performativity affects the reflexivity of data journalists in the devolved nations of the UK.

If journalism values and norms endured as the key element in the performativity of the data journalists interviewed during this research, they impacted their reflexivity to a similar extent. In spite of the

prevalent hybridity of their professional character, informants displayed a journalistic mindset and approach when discussing their work or specific cases. In this respect, an informant declared:

> 'For me, my mind might work slightly differently. I think about stories in terms of numbers. I think about what databases there might be behind it, and about the column and the rows, and how information is held. So I now think about my journalism differently, but still it's what is the news line, what's interesting here, what are the key facts, how does that relate to my readers' (DE1).

Whilst it is not a generalised trend in the devolved nations of the UK, data technologies have started to pervade the news production process of some professional newsrooms in the regions. A number of informants displayed some hesitancy about whether journalists should gradually adhere to computational logics, remarking that advanced computational thinking and knowledge should remain within the domain of programmers and computer scientists and did not necessarily concern journalists (DE2). Some informants observed that their know-how was perhaps rather elementary from a computing perspective (DJ1, DJ2, DJ6, DE1) but not necessarily any less powerful (DE1). One informant claimed that the reasoning behind remaining fairly simple in her computational analyses was that 'I don't want to over-corrupt it or overcomplicate it in a way that is unnecessary' (DE1). Amid a subtle aversion to adopting computational thinking, long-standing claims tend to emerge that data journalism is just plain journalism (DJ4), which seemingly constrains the development of more automated computer-mediated forms of data-driven journalism in the regions.

Within this reflexive dichotomy, and moving slightly away from the more traditional journalistic reflexivity, one group of informants did appreciate a progressive modification in the mindsets of data journalists. As one informant observed: 'We are really at the early days of something that is more than journalism, it could be art […] The way that stories can be told without words is huge' (DJ1). Despite a predisposition to praise the journalistic reflexivity of data journalism over computational reflexivity (DJ5, DJ6, DE1), informants referred to a new reflexivity mediated by the importance of understanding how coding works (DJ1, DJ2, DJ3), which would lead to a 'better appreciation of where they can find information and the tools to treat and interrogate that information' (DE2). The reflexive transformation mediated by the 'hybrid journo-coder' (DJ3) mindset appears to articulate a better understanding of when visualisations are required to enhance the story, a different mindset when

approaching stories (by seeking trends and patterns (DJ1), a different periodicity of news (that is driven by data releases instead of by breaking news), and an appreciation of the potential of automation (DJ6) and web scraping (DJ3, DE2) for news reporting. Although journalistic attitudes remain deeply ingrained within the reflexivity of data journalists in these devolved nations, a subtle appreciation of what a *techie* approach could bring to the trade is perceivable from the testimonies of informants.

Conclusions

This article has articulated a general epistemological framework of data journalism in Scotland, Northern Ireland and Wales using the prism of material, performative and reflexive conceptual lenses. Following the analysis, I conclude that data journalism in the devolved nations of the UK legitimises its knowledge claims by amplifying the existing distinctive epistemological character of the news media of these regions. With extraordinary potential for generating exclusive stories that normally remain hidden in the data infrastructures of local government, data journalism in Northern Ireland and Scotland has produced knowledge that copes with London-based and international competitors, appeals to local audiences, and makes use of locally-produced data managed through devolved policies. Although Wales-based data journalists were also able to generate exclusives, their knowledge claims appear to remain rather homogeneous, as they made use of national data that was repurposed and infused with local flavour by local reporters later in the process.

Knowledge production within 'national' outlets appears to be predominantly driven by legacy performativities and reflexivities. Across the three regions, there was generalised scepticism within legacy newspapers about the potential of data journalism. However, legacy organisations, such as the BBC Scotland and Trinity Mirror (Regionals) hired full-time staff to perform as data journalists, allocating more resources to this practice, whilst STV published sporadic pieces of online data journalism that were produced by staff in their spare time. Non-legacy organisations and actors were also deemed to be key in shaping the epistemological character of data journalism in the regions, with online investigative journalism outlet thedetail.tv in Northern Ireland and freelancers in Scotland being essential in the consolidation of a discreet but healthy hyperlocal data journalism ecosystem. Although other non-legacy actors, such as archivists, statisticians, civic activists, etc., were valuable

in shaping knowledge claims, their contribution in many cases remained tangential to the core of the news production process, which was predominantly managed by the data journalist.

When producing their knowledge, data journalists displayed a blend of journalistic and computational skills as they moved back and forth between the *newshound* and the *techie* approaches. In many cases their data—strongly intermediated by FOIA and open data regimes—was claimed as a transactional object that either informed their stories as evidentiary material or powered datafied outputs, thus restraining data's capacity for secondary agency. Computerised algorithms were developed and implemented by a minority of data journalists, leaving data journalists with no programming skills dependent on the constrained range of options offered by generic third-party software packages, which are limited by the universality of their design. In spite of the financial pressures that restrain local media in the regions, as data technologies pervade newsroom cultures, subtle traces of computational performativity and reflexivity emerge when practices such as web scraping, search for numeric trends, algorithmic data processing, or the generation of multi-layered datafied news outputs demand different methods of engagement from both data journalists and their audiences, consequently leading to their institutionalisation.

Based on the distinctions outlined above, it is possible to characterise data journalism in the devolved nations of the UK as a constant interplay between the distinctive identity of the devolved media and the emergent character of alien datafied philosophies. To produce their knowledge and to persuade regional audiences to accept their knowledge claims, data journalists engage in a continuous negotiation of power to advance regional hyperlocality against London-based and international competition, and counter precariousness within legacy organisations with resourceful non-legacy diversity.

Finally, this research encountered limitations when assessing the full extent to which data journalists' knowledge claims emerged from their adherence to one approach or the other, or, on the contrary, knowledge claims emerged from the data journalists' ability to switch back and forth between both approaches, depending on the task at hand. In this vein, future research would benefit from incorporating ethnographic observation as a methodological complement to the semi-structured interviews.

Acknowledgements

Notes

[1] Fink and Anderson (2015: 468) and De Maeyer et al. (2015: 439-40) have also remarked on the analytical challenges that the wide range of conceptualisations of data journalism poses for its scholarly study.

[2] The transfer of powers from the Parliament of the United Kingdom to the Scottish Parliament, the National Assembly for Wales, the Northern Ireland Assembly and the London Assembly.

[3] Computational thinking refers to an aspect of human cognition that draws on concepts from computer science and seeks to understand human behaviour through a problem-solving and system-design mentality (Gynnild, 2014: 723). Gynnild observes that the infusion of computational thinking into professional journalism challenges the 'fundamental thought system in journalism from descriptive storytelling to abstract reasoning, autonomous research and visualisation of quantitative facts' that equips journalists with 'complementary, logical and algorithmic skills, attitudes, and values' (2014: 725).

[4] Because it has its own public data records and issues, Scotland is not part of the unit's core remit and the data they tend to assess do not have Scottish relevance (DE2).

[5] For details, refer to Karlsen and Stavelin (2014: 41), De Maeyer et al (2015: 439), Fink and Anderson (2015: 477) and Parasie and Dagiral (2013: 858).

[6] For details, see Parasie and Dagiral (2013: 861-62), Fink and Anderson (2015: 470), De Maeyer et al (2015: 437), Karlsen and Stavelin (2014: 39), and De Maeyer et al (2015: 442).

[7] More information in Fink and Anderson (2015: 475), Appelgren and Nygren (2014: 401), De Maeyer et al (2015: 441).

References

Anderson CW (2013) Towards a sociology of computational and algorithmic journalism. *New media & society* 15(7): 1005–1021

Anderson CW (2015) Between the Unique and the Pattern. *Digital Journalism* 3(3): 349–363

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

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

Broersma M (2010) Journalism as performative discourse. The importance of form and style in journalism. In: Rupar V (ed) *Journalism and meaning-making: Reading the newspaper*. Cresskill: Hampton Press, pp. 15–35

Coddington M (2015) Clarifying journalism's quantitative turn. *Digital Journalism* 3(3): 331–348

Conboy M (2011) *Journalism in Britain: a historical introduction*. London: Sage.

Coole D and Frost S (2010) Introducing the New Materialisms. In: Coole D and Frost S (Eds) *New materialisms: Ontology, agency, and politics.* USA: Duke University Press, pp. 1–43

Cushion S, Lewis J, and Groves C (2009) Reflecting the four nations? An analysis of reporting devolution on UK network news media. *Journalism studies*, *10*(5): 655–671

De Maeyer J, Libert M, Domingo D, Heinderyckx F and Le Cam F (2015) Waiting for Data Journalism. *Digital Journalism* 3(3): 432–446

Dekavalla M (2015) The Scottish newspaper industry in the digital era. *Media, Culture & Society*, *37*(1): 107-114.

Fink K and Anderson CW (2015) Data Journalism in the United States: Beyond the 'usual suspects'. *Journalism Studies* 16(4): 467–481

Guillaume X (2002) Reflexivity and subjectivity: A dialogical perspective for and on international relations theory. *Qualitative Social Research* 3(3). Available at: http://www.qualitative-research.net/index.php/fqs/article/view/826/1795

Gynnild A (2014) Journalism innovation leads to innovation journalism: The impact of computational exploration on changing mindsets. *Journalism* 15(6): 713–730

Hamilton J and Turner F (2009) Accountability through Algorithm: Developing the Field of Computational Journalism. Report from Center for Advanced Study in the Behavioural Sciences, Summer Workshop 27-31 July. Available at:

http://web.stanford.edu/~fturner/Hamilton%20Turner%20Acc%20by%20Alg%20Final.pdf

Jones A (1993) *Press, Politics and Society: a history of journalism in Wales*. Cardiff: University of Wales Press.

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

Lewis SC and Westlund O (2015) Big Data and Journalism. *Digital Journalism* 3(3): 447–466

McNair B (2006) News from a small country: the media in Scotland. In: Franklin B and Murphy D (Eds.) *Local Journalism and Local Media: making the local news*. London: Routledge.

McNair B, Dekavalla M, Boyle R and Meikle G (2010). Mapping futures for news: trends, opportunities and challenges for Scotland. *http://www. scottishinsight. ac. uk/portals/50/futures_news_report. pdf*, 1-39.

Miller D (2005) Materiality: an introduction. In: Miller D (ed) *Materiality*. USA: Duke University Press, pp. 1–50

Parasie S and Dagiral E (2013) Data-driven journalism and the public good: 'Computer-assisted-reporters' and 'programmer-journalists' in Chicago. *New media & society* 15(6): 853–871

Parasie S (2015) Data-Driven Revelation? *Digital Journalism* 3(3): 364–380

Rolston B (Ed) (1991) *The Media and Northern Ireland: covering the troubles*. Basingstoke: MacMillan.

Schlesinger P (1998) Scottish devolution and the media. *The Political Quarterly*, 69(B): 55–74.

Schlesinger P, Miller D and Dinan W (2001) *Open Scotland? Journalists, spin doctors and lobbyists*. Edinburgh: Edinburgh University Press.

Tabary C, Provost AM and Trottier A (2015) Data journalism's actors, practices and skills: A case study from Quebec. *Journalism* 17(1): 66–84.

Van der Tuin I and Dolphijn R (2012) *New materialism: Interviews & cartographies*. USA: Open Humanities Press.