Pre-Print version

Porlezza, Colin, and Sergio Splendore. 2019. 'From Open Journalism to Closed Data: Data Journalism in Italy'. *Digital Journalism*, September, 1–23. https://doi.org/10.1080/21670811.2019.1657778.

From open journalism to closed data
Data journalism in Italy

Colin Porlezza (corresponding author)
Department of Journalism
City, University of London
United Kingdom
E-Mail: colin.porlezza@city.ac.uk

Sergio Splendore
Department of Social and Political Sciences
Università degli Studi di Milano
Italy

Abstract
The study investigates how journalists construct the rise of data journalism discursively *within the specific political and economic conditions of Italy, and whether data journalists describe it as able to promote transparency at the level of both journalism and society as a whole. Italy is an interesting case for four different reasons: 1) to date, few analyses have been carried out on Mediterranean countries; 2) the Italian media system and its traditional journalism culture appear to be reluctant to innovate; 3) Italian data journalists are generally freelancers linked to small news agencies rather than established newsrooms; 4) the research reported in this paper was accomplished while the Italian government was approving a Freedom of Information Act. From the end of 2015 to the beginning of 2017, we carried out 15 interviews with data journalists working full-time both for established news organizations and specialized agencies. The findings show that data journalism in Italy is a highly professional sub-field, although journalistic education is not well developed. Data journalism is still largely determined by the availability and accessibility of public datasets. However, data journalists have developed certain strategies with which to generate and collect their own data, for instance collaborations and networks.*

KEYWORDS
data journalism; datafication, FOIA, journalism education; open data; public administration; transparency

Italian journalism culture meets data journalism

The data turn in journalism, or the 'interactive turn' as Usher (2016) would term it, has shown, both through newsroom experience and media research, that practices relating to the field of data journalism may yield several advantages: accessing datasets and neglected stories hidden in public data; fostering democratic conversation with the audience (Boyles and Meyer 2016); improving investigative reporting (Flew et al. 2010), more transparency (Stoneman 2015) and a newly reinforced watchdog role (Felle 2015). A rapidly growing body of research proves that data journalism is on the rise in many different countries around the globe, not just in Europe and North America, but also in Asia (Zang and Feng 2018) and South America (see Valencia and Restrep 2019). While there are no systematic studies on data journalism with regard to the African continent, Cheruiyot and Ferrer-Conill (2018) nevertheless show that there is an increasing push for data-driven strategies in the news media as well. However, "contemporary forms of use of data in journalism production are multifaceted" (Splendore 2016a) and unevenly distributed across countries and media segments, since journalistic practices and economic resources differ significantly among news organizations. Coddington confirms this view when he concludes that the use of data "has splintered into a set of ambiguously related practices" (2014, 332).

In this paper, as our theoretical framework we apply a sociological perspective to study Italian data journalism. According to Shoemaker and Reese's (2014) hierarchy of influences, there are five categories to investigate determinants of media content: social systems, social institutions, media organizations, routine practices, and individuals. In describing data journalism in Italy we will try to investigate at least four out of these five levels. As Anderson (2013) states, applying a sociological perspective to study data journalism means first that the enactment of the various forms of data journalism depends on social institutions that lie outside journalism itself (Anderson 2013: 1009). Therefore, we analyze data journalism primarily on an inter-institutional level, where the journalists' work and performances depend on relations with other institutions such as, for instance, the government or political parties. We also consider context and media organizations where data journalists work for. And we also analyse journalists' practices and routines. As Ekström and Westlund (2019: 264) affirm, applying this perspective to journalists' epistemology, a sociological approach understands the processing of facts (data in our case) as practical matters handled through norms and standards developed within a particular context. That is precisely what we do in analyzing Italian data journalism by taking into account the context in which it emerges, and the norms (in particularly transparency) data journalists enact. Finally we will consider the individual level, especially regarding Italian data journalists' education.

In our approach, considering the political context means looking at the broader "[d]evelopments in the world of so-called 'open-government initiatives,' the role of large-scale databases in the crafting of public policy, the transparency and accessibility of government data and its use in democratic decision-making" (Anderson 2013, 1012). What Lewis and Al Nashmi (2019) would call structural barriers to the development of data journalism. As Appelgren and Salaverría (2018) maintain in a study on data journalistic practices in Sweden and Spain, it is important to analyze data journalism development in light of the countries' approaches to public transparency.

From an economic perspective, we mainly consider how the Italian media system in general and newsrooms in particular have triggered and encouraged forms of data journalism. In Shoemaker and Reese's (2014) terms, the political environment represents here the institutional level, while the economic one also the media organization level. Additionally, we consider also how and whether quantitative forms of journalism fit general attributes of the Italian

journalistic culture. Famously, the Italian journalism culture is characterized by the Mediterranean countries' general features: a high level of political parallelism, a journalism more strongly oriented to comments, and a low level of professionalism (Cornia 2014; Hallin and Mancini 2004). In this paper, we do not focus exclusively on the three main conditions (political, cultural and economic) that characterize the Italian media system, but we apply a broader approach, which is why we do need to take them into account when conducting an overall assessment of data journalism in Italy. Hence, this paper continues with the tradition of those initial country-specific data journalism studies carried out in Belgium (De Mayer et al. 2014), Norway (Karlsen and Stavelin 2013), Sweden (Appelgren and Nygren 2014), and the United States (Fink and Anderson 2015) that applied an early explorative analysis to an emerging field, focusing on "fractures, fissures, and power-dynamics at work within that field, as well as the way that this field is shaped by other institutional clusters in adjacent spaces" (Fink and Anderson 2015, 468). Particularly in a country like Italy, where the interrelations between politics and journalism are so close - although especially the press benefits from a growing autonomy due to an increasing commercialization and a strong adversarial coverage (D'Arma 2015) - looking at the structural level of interfering institutions is essential.

We also investigate how data journalists understand and implement the norms of transparency and accountability. Previous research on data journalism in other countries has reported similar statements by data journalists, criticizing the dominant journalistic culture as "lacking a culture of transparency" (De Mayer et al. 2015: 442) - even if transparency is generally a value that journalism strives for (Singer 2007). However, also in the case of data journalism, the transparency norm - understood here as disclosure transparency (Karlsson 2010), which means that journalists ought to implement openness regarding the production, selection, and analysis of news (Tandoc and Oh 2018) and the sharing of open source code accompanied by relevant documentation (Lewis and Usher 2013: 607) - is not always implemented: some studies have shown that transparency in terms of access to the underlying data is provided in most stories (Young, Hermida, and Fulda 2018; Tandoc and Oh 2017; Parasie and Dagiral 2012; Loosen et al. 2017), but other findings relativize data journalism's openness (Lowrey and Hou 2018; Porlezza 2019; Zamith 2019). In synthesis: data journalism depends on the specific structural and institutional framework in a given media system - which is the main reason why there are still so many differences both within (see Borges-Rey 2017) and between media systems in relation to journalistic practices and performances if one looks beyond the usual benchmarks (Splendore 2016b).

The research reported in this article consisted of 15 semi-structured interviews conducted with Italian data journalists while the Italian government was debating the implementation of a Freedom of Information Act law. In Italy, the Council of Ministers approved the so called "transparency decree" on 25 May, 2016. It was followed by the so-called *decreti attuativi* (the rules by which a decree is applied). The interviews were thus carried out in a very specific moment useful to understand and interpret data journalists' role perceptions, but also their reflections - elicited by the discussion with respect to the FOIA - regarding the changing political context, and thus the changing structural and institutional context in which data journalism has arisen in recent years. We thus study the practices and role perceptions of data journalists in Italy according to their discourses about the rise of data journalism in both legacy news organizations and specialized news agencies. Two research questions are therefore addressed: First, what kind of data journalism do Italian data journalists conduct within the specific political and economic conditions of Italy? Second, do data journalists aim to promote transparency both at the journalistic level and at the level of society as a whole? In line with De Maeyer et al. (2015), the paper's purpose is therefore to analyze how journalists construct data journalism discursively, since their positions are crucial in order to conceptualize "the modes of existence of this emerging news practice" (De Mayer et al. 2015: 435). The analysis

is thus carried out within the framework outlined above: the economic and political context, the journalistic culture and the norm of transparency, and the function of criticism and control in light of the newly introduced FOIA. Besides the fact that the research was conducted while the Italian government was approving a Freedom of Information Act, Italy is an interesting case for three different reasons: 1) to date few analyses have been conducted on Mediterranean countries (see Appelgren and Salaverría 2018), and none for the Italian context; 2) the Italian media system and its traditional journalism culture appear to be reluctant to innovate (see Agostini 2004; Splendore 2017); and 3) Italian data journalists are generally freelancers linked to small news agencies rather than traditional newsrooms. The article has four sections: the first critically discusses previous research on data journalism and its definition(s). The second section provides a description of our methodology, while the third section presents the results of our study. The fourth section instead conducts critical assessment of the findings and discusses further implications for both Italian data journalism and data journalism research. The article concludes with a discussion of the study's limitations and suggests matters for future research.

Literature review and the Italian context

Kennedy et al. (2015) describe datafication as "the process of rendering into data aspects of the world not previously quantified" (Kennedy et al. 2015: 1). This process occurs not only in society at large; it is also reflected in journalism as well, particularly with regard to quantitatively oriented forms of journalism (Porlezza 2018). Data journalism is one of the phenomena emerging in the news media that are directly related to the increasing datafication of society. Although the amount of scientific literature on data-driven forms of journalism (Coddington 2014) is steadily increasing, there is still no commonly shared definition: various labels such as 'data journalism', 'computational journalism', 'programmer journalism', 'data-intensive newswork', or even 'robotic' or 'algorithmic' journalism are in use. Detailed typologies are discussed by Anderson (2015), Coddington (2019), Ferrer-Conill (2017), Splendore (2016a) and Usher (2016). From a scholarly point of view, categorization is important because it reminds us that these terms and practices have institutional, cultural, and epistemological roots (Coddington 2019: 226). Ausserhofer et al. (2017, 3) identify definitions that focus on the change in the process of newswork from those which emphasize that data journalism produces news items based on data analysis and interactive visualization.

The latter is the starting point suggested by Usher (2016), who provides a clear overview of the different labels. In her research, accomplished through several interviews and ethnographies conducted within newsrooms around the globe (*The Guardian*, *The Washington Post*, *FiveThirtyEight*, *AlJazeera*, and others), she establishes two unavoidable borders that define the various forms of data-driven journalism: First, those labels depend mostly on the kind of education received by practitioners. Essentially, she maintains that those who come from a journalistic background tend to define the work as data *journalism*. Instead, those who have a programmer or computer science background tend to define it as *computational* journalism. In Italy, data journalism is carried out mainly by journalists (who have therefore received traditional training) rather than programmers, who put their expertise at the service of journalism (Splendore et al. 2016). In Italy, "hacker journalists" are therefore rather the exception than the rule, although there is a certain blending between journalism and computer science: similarly to what Lewis and Usher (2014) observed in the U.S., there are Hacks/Hackers chapters in Italy as well. However, it is more common for journalists to try to acquire computational expertise rather than the other way around, that is for computer specialists or programmers to become experienced journalists.

Nevertheless, research shows that in times of datafication, journalists and other social actors are increasingly sharing similar skills and aspirations. Baack (2018) for instance shows that data journalists and civic technologists are acting in complementary ways, creating communities of practice of *flexible data professionals* who "aspire to work in a public interest: they share transferable skills in dealing with data and using information rights, and they want to use these skills to create public services that hold powerful people and institutions accountable and empower citizens" (Baack, 2018, 688). Similarly, Ananny and Crawford (2015) observed the creation of new media actors.

However, the complementarity of activities are not only related to the community of civic technologists. Data journalists also engage in forms of *inter-organizational networked journalism* as described by Hermida and Young (2017), where journalists are collaborating with actors from outside the established field of journalism. This might be, as the authors state (Hermida and Young 2017, 174) "predictive of a more itinerant and unfixed professional workplace context and identity". The creation of new interdisciplinary and collaborating communities such as Hacks/Hackers, or the formation of entirely new communities of practices of flexible data professionals might just confirm the authors' prediction.

As other actors establish closer ties with journalism, they start to exert effects both on workflows and, by doing so, can also drive innovation in the news production (Belair-Gagnon and Holton 2018), becoming *agents of change* (Eldridge II 2018). It is however important to note that such collaborations evolve and emerge over time given that initiatives to foster more open and participatory initiatives are initially met with hesitancy (Belair-Gagnon and Holton, 500). Generally, collaborations between journalists and other social actors are becoming more common, which is also reflected by the increasing scholarly research that investigates these phenomena (see also Deuze and Witschge 2017, Lewis 2012; Lewis and Usher 2014).

Secondly, Usher (2016) considers also the importance of how practitioners themselves define their work. Accordingly, she concludes that "Europeans call data-journalism what Americans define computational or programmer journalism" (2016, 38). This is the reason why we adopt the label *data journalism* in this analysis of the Italian case, because Italian journalists use that label. Nevertheless, following Coddington (2019), the fact that Italian journalists call it 'data journalism' means also, from an institutional perspective, that they demand transparency from the public administration, and from an epistemological perspective they regard data as *the underlying evidence* of their news items. Italian journalists define data journalism as an expression of different forms of data-driven products generated by various processes that clearly differ from traditional news production. In their view, the use of this label therefore goes well beyond the simple and static forms of infographics, mostly published on paper.

This study adds to the many single-country or single-case studies. It is an exploratory work that tries to fill a gap due to the absence of Mediterranean countries, and, specifically, of the Italian case, among the studies already accomplished (see among others Appelgren and Nygren 2014; Appelgren and Salaverría 2018; Fink and Anderson 2014; Borges-Rey 2016, 2017; Boyles and Meyer 2016; De Maeyer et al. 2015; Karlsen and Stavelin 2013; Lewis and Usher 2014; Paraise and Dagiral 2012; Tabary et al. 2016; Weinacht and Spiller 2014).

However, to study data journalism in Italy means that it is necessary to go beyond the traditional newsrooms (Fink and Anderson 2014) and to look at other actors in the field, particularly those at its border. It is more likely to find data journalists outside established newsrooms. This makes the Italian case different from that of other analyzed countries such as Sweden (Appelgren and Nygren 2014), the UK (Borges-Rey 2016) or cities like Chicago (Paraise and Dagiral 2012), also because there are no other cases where freelancers are more important than those working in legacy news media. This aspect is increasingly important for the development of data journalism, as showed by Cheruiyot an colleagues (2019) who have investigated the role of a diverse set of actors, which contribute to data journalism outside

legacy media. The uniqueness of the Italian case - both in terms of the structure of its media system and the origins of data journalism, make a strong case for its deeper analysis.

While the dominance of data journalism freelancers is an Italian specificity, the other issue that we investigate here is common to several studies in data journalism: the relationship between transparency and journalism. The fact that data journalism as a political intermediary can enhance the knowledge and use of openly available government data and thus promote transparency and participation is almost commonsensical (Brugger et al. 2016; Felle 2015). By using data, journalists perform the democratic duty of helping citizens to engage with publicly available datasets and to become more actively involved in the process of their interpretation (Boyles and Meyer 2016). However, even if this process seems to be straightforward at a theoretical level, the potential impact of data journalism has to be studied within the political, social and cultural contexts that shape its forms and its understanding (Anderson and De Maeyer 2015). While we cannot investigate the effective impact of data journalism, we are still able to analyze the journalists' accounts with regard to the limitations and the power of their work.

In their everyday practices, data journalists constantly struggle with technologies, logics and dynamics that either enhance or undermine their performance (Lewis and Usher 2014). A major issue for data journalists is access to data: Borges-Rey (2016), for instance, shows how data journalists in the UK struggle when it comes to holding data organizations and brokers to account because their ability to access data, compared to the public administration bound to the FOIA, is limited. In this regard, the availability of publicly accessible data has, in fact, a direct impact on the journalists' ability to carry out their primary function of control (no access to data = no control = no transparency). This is also a difficulty related to the Italian environment, even if the political and cultural contexts in Italy are far more challenging (Porlezza 2016): The longstanding non-existence of a FOIA, and the limited availability of open data seriously hampered journalists' work in Italy. Generally, Italian public data providers are less likely and generally less equipped to provide open data ready for analysis by journalists.[1] Moreover, the lack of a widespread data literacy also prevents investments by media outlets in forms of data-driven news reporting. An open data culture is, as several studies have shown (Aitamurto et al. 2011; Appelgren and Nygren 2014; Baack 2015; Sandovàl-Martin and La-Rosa 2018), a prerequisite for a flourishing data journalism. The late establishment of an open data culture in Italy, particularly in politics and the public administration, has inhibited the development of data journalism. In Italy in particular, the open data culture (and movement) contributed to awareness about the importance of open data, but eventually it led to the creation of the FOIA and, therefore, allowed data journalism to reach another level (Porlezza 2016).

Before proceeding with the discussion of results, it appears necessary to dwell briefly on contemporary Italian journalism. We have already mentioned that Italian journalism culture is commonly depicted by applying the widely-known Mediterranean model to the Italian context (Hallin and Mancini 2004) and how it differs from other journalistic cultures in Europe (Hanitzsch et al. 2011). Along these lines, the recent publications from the Reuters Institute for the Study of Journalism at Oxford highlight how forms of politicization of the Italian news ecosystem are still evident and how publishers encounter severe difficulties in finding new news products for the Italian market given that, for instance, newspaper readership continues to fall (Reuters 2019). In this regard, particularly the Italian print press is characterized by an elitist dimension, which is also demonstrated by the low levels of readership, and by a more advocate and interventionist stance - sometimes explicitly rejecting journalistic principles such as objectivity (Mancini 2005). Moreover, Italian journalism is usually regarded as historically reluctant to innovate (Agostini 2004). Nevertheless, digital disruption is bringing some innovations, or at least some forms of hybridization. As Bentivegna and Marchetti (2018)

highlight in their study on Italian journalists' use of Twitter, evident is "a restatement of traditional journalistic practices combined with the adoption – or sometimes the mere exhibition – of the norms and practices typical of the participatory web". Also Splendore (2017) shows that in terms of sources there is an increasing mix between institutional and non-institutional ones.

Methodology

The paper is based on 15 semi-structured interviews conducted with selected data journalists from different organizations such as traditional news outlets, journalism start-ups and independent journalistic platforms. This method has been employed in several other studies on the emerging data journalism landscape (Fink and Anderson 2014; Parasie and Dagiral 2012) precisely because it investigates properly new facets of a phenomenon. The first interviewees were handpicked on the basis of their involvement in data journalism practices and this was verified when he or she was first approached. Subsequently, the authors adopted the strategy of "snowball sampling" (Becker 1963). In this method, the researcher selects a couple of interviewees to start with and then asks them to nominate other potential interviewees (in this case other data journalists) relevant to the study. This kind of sampling method is particularly useful for analyzing highly networked communities such as data journalists (Young and Hermida 2017)[3] and it has been used by other researchers in the subfield of data journalism (see Felle 2015; Weihnacht and Spiller 2014) as well. Hence it is particularly suitable for investigating data journalism in Italy.

(Insert table 1 here)

The selection of new interviewees followed the concept of saturation: we stopped our interviews when they were not telling us anything we had not heard before (Edwards and Holland 2013). The first interviewee contacted was a member of the Venice chapter of Hacks/Hackers (see Table 1 for an overview of the most important agencies that work in the Italian data journalism context). Apart from being active in data journalism, the selected person was also chosen because of his/her? affiliation to Hacks/Hackers, which represents an example of an innovative and hybrid institution positioned between journalism and programming (Lewis and Usher 2014). Consequently, the authors contacted also different people named in this first interview. The fact that Italian data journalists are closely interconnected and networked was also confirmed by several names being mentioned more than once. Already at this stage of the project it seemed that the data journalism scene in Italy was rather small but extremely well connected. Among the interviewees there were journalists working in both traditional news organizations and journalistic news agencies specialized in data journalism.
Overall, we interviewed 15 journalists, five of whom were female. On average their age was between 30 and 40 years old. None of the interviewees had attended a proper school of journalism, but many had spent periods of internship within well-established newsrooms. However, almost all of the data journalists came from a journalistic background. Except for a programmer, all others had acquired their coding expertise either by attending courses offered by international organizations such as the European Journalism Centre or the Centre for Investigative Journalism – both of which are non-Italian organizations. Comparing both the background and the educational paths of the interviewees results in a very heterogeneous sample, but this is quite common among data journalists, as previous research has shown (Splendore et al. 2016).
Obviously, this kind of method has some limitations. The sample was by no means representative, and we could also have been confronted with citation cartels in such a closely-

connected network. However, given that we were not only offered the names of well-known data journalists, and after 15 interviews no new candidates were mentioned, it does not seem that we were dealing with a citation cartel. Snowball sampling is debatable, but in this specific circumstance it made sense to question data journalists - the experts themselves - about their peers because there is still no common definition in journalism research, which makes the description as well as the demarcation of what is data journalism and what is not extremely difficult and epistemologically challenging. It is not possible to determine an exact population of data journalists, particularly outside the established newsrooms, but considering the names repeatedly mentioned in the interviews, the sample might well have reflected the small and active data journalism population in Italy at the time of the study.

The interview guide was structured into four main areas: a) practices and careers; b) data journalism, its definitions, functions and impact in the Italian public sphere; c) the roles covered by data journalists within their organizations; d) professional and ethical issues with regard to data journalism. Each interview included also a "reconstructive" part when journalists were asked to explain and evaluate the practices involved in the journalistic production process based on a specific article or story produced by them. We then asked them to reconstruct their work in terms of sources, data analysis and visualization, journalistic transparency, and finally with regard to their journalistic aims. This method has demonstrated its viability in exploring different facets of news processes (see Godler and Reich 2017). Reconstruction interviews, especially when aimed at investigating sources, are usually conducted in special settings used to avoid infringement of source confidentiality. We opted to include this part in the flow of the semi-structured interviews, since our goal was to understand what influences journalists regard as important in their work and also what aims they pursue in their daily production of data journalism items. Additionally, reconstruction interviews were designed particularly to investigate the second research question, i.e. whether and how data journalists provide transparency in what they produce.

The interviews lasted between 50 and 70 minutes and were carried out by means of Skype, thus following a methodology successfully used in previous studies (Hanna, 2012). The interviews were transcribed and then cross-analyzed. From areas a and c (see above) we extracted the information elaborated in the subsections "Political and economic influence perceived by Italian data journalists*"* and "Opening up institutional archives*".* From area d we derived the information elaborated in the subsection "Open data journalism and transparency". And from area b we generated the information regarding the definition of data journalism (already discussed above). The interviews were carried out between the end of 2015 and 2017. All interviewees were assured anonymity.

So close(d): opening Italian data

In order to proceed with discussing our main research questions, it is necessary briefly to describe the data journalism field in Italy. This brief overview is mainly based upon data journalists' statements. Data journalism in Italy is still a new and innovative area of specialization within journalism. The slow development of data journalism in Italy can be related to the distinctive Italian journalistic context, which historically has been laggardly in adopting new and innovative procedures - sometimes even refusing them (see Agostini 2004; Splendore 2017). Some of our interviewees estimated the number of data journalists working full-time in Italy at around 20, most of whom are freelancers. The number certainly increases if journalists who deal *also* (but not mainly) with data journalism are included. Although the number of traditional media outlets presenting some data journalism items on their websites is increasing (see *Espresso*, *AGI*, *Corriere della Sera*, *La Stampa*), almost all journalists working

in the field of data journalism are affiliated to one or more specific data journalism agencies such as Dataninja or Datajournalism.it, or the emerging Visual Agency. Data journalists in Italy form a highly interrelated network organized more as an emergent network than as a scattered group of specialists located in different news outlets across the country. As we shall see, the network is very collaborative in diverse areas.

Political and economic influence perceived by Italian data journalists.

In regard to influence from a broad political perspective, three different themes were highlighted by our interviewees. The one most often mentioned was the difference between data and traditional journalists. Italian data journalists very often stressed their independence and autonomy compared to the traditional journalists, who were regarded as being highly influenced by politics, parties and politicians. Although important, this discourse boils down to their professional values rather than the effective influence of the political context.
The other two recurrent issues concerned the approval of a FOIA law and informal collaboration with public servants. As regards the FOIA, the commonly shared hope was for "a game changer". As one of the journalists told us: "A FOIA would be the best. That way we could access gold mines from their archives" (j8).[2] However, the network that includes people related to data journalism also comprises public servants who help data journalists to find and gather useful data and information.

The Italian public administration lacks a *data culture* [e.g. it gathers and provides just a few data to analysis], sometimes not revealing its mistakes on purpose […]. At the same time, several smart and competent people, like Spaghetti Open data, which are part of our network, try to help us, but they are not [i.e. within the public administration for which they work] those in charge of deciding (j8).

This means that there is close collaboration between data journalists and some civil servants. However, the interviewee specifically refers to the existence of a "data culture" that is often missing in the public administration, but very much present in the data journalism community. This "data culture" is usually defined in terms of a new climate of collaboration (Porlezza, 2017), innovation and creativity (Young and Hermida 2017), but also in terms of data literacy - simply by being able to understand the importance of data for the public interest. Some of these are specialized actors in the field such as e.g. Spaghettiopendata, a network of investigative and data journalists able to gather and generate data independently from public sources.

Sometimes I ask other persons involved in our community for help, for instance Spaghettiopendata or other groups. Usually a close collaboration exists, we help each other, even when we work for different media outlets (j2).

This kind of collaboration is aimed at making the public administration more transparent. Collaborations between data journalists and other social actors such as open data activists, hackers or hacktivists occur increasingly often – and are therefore also of growing academic interest. However, the interpretation about the political influence on data journalists' work is completely focused on the availability and accessibility of data from the public administration, rather than politicians and politics. In this regard, during the reconstruction part of the interviews, interviewees told several anecdotes in accounting their practices for the production of news.

Getting those data from our Ministry was a nightmare! When I was lucky enough to contact them via telephone, they agreed to send me the data via mail! The sent to me the printed book with plenty of tables. I finally and

luckily found another open dataset online with those data, but provided by an NGO. Their book was useful to verify the veracity of the data I found online (j6).

Regarding the economic influence, the interviewees' discourses mainly centered on the fact that the richest and most established newsrooms do not invest in data journalism. The development of Italian data journalism is therefore limited due to the lack of investments by the publishers. At the same time, the fact that the field remains very small generally helps data journalists to find different ways to earn a living, usually by mixing their work as data journalists with other activities – often in the field of activism related to NGOs.

The problem is that everyone [i.e.: every Italian data journalist] follows different projects, that's why we can make data journalism full-time (j12).

Data journalism opened up different perspectives […] just a few people produce data journalism, there is a lot to do and there are a lot of stories to tell by the use of data, we do not lack ideas (j5)

Nonetheless, data journalists complain that they cannot undertake more complex projects. The problem is the absence of foundations or (media) organizations that are eager to invest more in data journalism. Nevertheless, this situation has increased their expertise in competing and not rarely gaining grants from international organizations.

I also run projects that last for six months, but these projects need external grants, not just journalistic ones. Foundations, NGOs, on different occasions we obtained a grant from organizations like journalismfund.eu or the European Journalism Center. (j8)

Open data journalism and transparency

The network of data journalists in Italy uses different platforms to share data, practices, opinions and suggestions on how to analyze data, tools to use etc. on their own. Consequently, in such an environment, journalists also share values and their conception of data journalism itself. Unlike journalists educated and socialized in traditional newsrooms in Italy, data journalists share professional norms and values that are strongly related to an Anglo-American culture of journalism (Kovach and Rosenstiel 2014).

Data journalism basically originates from Anglo-American journalism […] There, journalism is much more oriented towards data and facts, not towards opinions like in Italy. Obviously, if a journalist talks about a specific topic, it should be important for her, but I don't think that my pieces are automatically opinionated (j6).

The interviewees pointed out in regard to the storytelling part of their work that one of their main concerns was to rely as much as possible on the data that they had at hand and to avoid any kind of personal or opinionated statement. This professional conception was confirmed by the strong commitment to the notion of transparency. Many data journalists agree that if (raw) data are not shared and made public, journalists become part of the same system that tries to keep things secret: "News organizations may be campaigners for open information but by withholding that data, become complicit in a system which essentially keeps data private until it's no longer commercially valuable. It's all very well calling for governments to throw open the doors of their data vaults, but if you are not willing to be open too, what is that worth?" (Rogers, 2012). Therefore, transparency has two different meanings: First, it is seen as a goal to achieve when it comes to the government or the public administration (see below). Second, for Italian data journalists transparency means also being transparent about their own work in order to be held to account. In this sense, the concept of "open journalism" becomes a central value. It means sharing datasets among different journalists. It also means that other actors such

as NGOs and the wider public are able to access and use the data for further collaboration or investigation.

I absolutely agree to the idea of open journalism. It would be a paradigmatic change in journalism if the production processes were open and transparent. (J3)

According to me, open journalism is one of the keys that allow people to understand how journalistic labor and mediation can make sense of openness and the availability of data. (J2)

Open journalism becomes not only an approach useful for making the journalistic production process more transparent. It also serves as key to fundamental normative changes within the (data) journalistic profession by seeking to establish the open-source notion of sharing and networked journalism as a new standard (Porlezza 2019).
In general, all forms of journalistic transparency are taken for granted by data journalists in their practices and routines. For instance, they agree to show a methodology that explains where the data have been collected. They also expect to see some explanation with regard to the analyses that have been carried out, and they concur in conducting further controls to test the reliability of the databases used.

I'm very inspired by the Anglo-Saxon model, only a few frills, but data, facts! It is important to publish a methodological note. This gains points in terms of credibility and accountability because the reader can test the methodology (j7).

During the analysis, you can still introduce checks. If you foresee a series of problems, in the scraping phase you conduct a series of checks. […] If you do not avoid errors right from the beginning, they will explode (j11).

What is relevant is that the projects we have selected for the "reconstructive" part of the interviews (see the methods section) usually include those elements.[4] In our research design, the reconstruction interviews were particularly apt to reveal this attitude. In that moment of the interview, our interviewees were prompt in describing the parts of their work aimed at restoring transparency.

I usually provide a methodological note where I explain every step I have made during the collection and analysis of my data. When I teach beginners, I always repeat that this is the most important part of their project, even if readers do not care (j11).

Even though no one asks me, I always write a post for my medium.com account for all my news. (j4)

Our interviewees usually placed a certain degree of emphasis on the importance of that part of their news items.

Opening up institutional archives

If one of the most often shared values is transparency, another frequently mentioned normative principle of data journalists' work in Italy is *critique and control*. This function is closely in line with Felle's (2015) as well as Boyles and Meyer's findings. Their interviewees revealed that journalists directed their efforts to making "the information 'less clunky' than governmental interfaces" (Boyles and Meyer, 2016, p. 5). Many data journalists see themselves as an additional resource to the fourth estate – as watchdogs on those in power or, in less radical terms, as allies in making the public administration more transparent.
Data journalism can devise instruments and methodologies even more sophisticated than traditional investigative reporting and is thus capable of analyzing large datasets. This allows

data journalism not only to recombine already published datasets, but also to discover currently unknown wrongdoings. As mentioned above, this endeavor is strengthened by another distinctive characteristic of the network that operates around data journalism. It includes actual and aspiring journalists, programmers, but also several interest groups with whom people interested in data journalism gather and discuss current issues. In this environment, journalists' aims are mixed with those of civil servants to promote and increase transparency.

The main function of data journalism is to be a better watchdog. It can be a better supervisor of those in power, of the use of financial resources – but only if you are able to read and understand the data. For example, if you are able to read a certain budget and actually understand it, then you are a far better watchdog because you don't limit yourself to the publication of a press release issued by the public authorities (J5).

The importance and the result of making data publicly accessible is that data journalists enable other actors such as NGOs, programmers or hackers to further investigate and disseminate either the data themselves or the findings (Lewis and Usher 2013; 2014). Nevertheless, most journalists raise doubts about data journalism's ability to improve transparency in other institutions such as the public administration in terms of open data and open government: that is, to "increase the availability of information about governmental activities" to "support civic participation", to "implement the highest standards of professional integrity throughout our administrations" and to "increase access to new technologies for openness and accountability".[5] The interviewees generally agreed that the impact of data journalism on actors such as the public administration depends mainly on the enactment of a Freedom of Information Act.

At the moment, the impact of data journalism on the public administration is close to zero. It is possible that you can get in touch with a public official that is eager to help you – but in most cases, it is a bargain, also because the journalistic enquiry is not seen as something valuable. (J3)

The crucial point is that the public administration itself should understand openness and open government as an opportunity. As an opportunity to innovate processes and to make its work more efficient. As an opportunity to get closer to the citizens by opening up their data. As long as the public administration does not understand the positive effects of such a cultural change, data journalism alone will not be able to trigger this process of transformation (J2).

Since 2017, the most important data journalism projects have used the new FOIA. However, the opportunities provided by the new FOIA have been integrated with other strategies that data journalists already applied. For instance, FOIA requests are made even though journalists, at the same time, try to gather data autonomously – as they have done, for instance, in the investigative reportage entitled "The Migrant Files", winner of the data journalism award in 2014. Or else they rely on data produced by other institutions such as NGOs.
There are different elements at play to open the public administration while reinforcing democracy (Felle 2015). It still remains to be seen (i) how the new FOIA, established only in May 2016, will be actually applied in Italy; (ii) whether and how the open data culture will spread, and whether public servants will actually allow access to important information and, eventually, collaborate with journalists – maybe also through the use of whistleblowers; (iii) the ability of journalists to minimize the risks of using alternative sources.

I usually gather my data from various associations, but it is obvious that in general data journalists work with open data from the public administration. But it would be more important to gain reliable and private datasets, rather than those to which the public administration allows us access (j14).

In an environment where data are not commonly open, relying only on datasets provided by the public administration – that at the same time is becoming apt to use open data at its

convenience – could undermine the role of watchdog performed by Italian data journalism. Data journalists themselves are aware of this risk, since they distinguish between more ambitious projects that require the collection and generation of data from scratch, and those projects that simply rely on publicly available data from the public administration.
Overall, the interviewees consistently referred to the issues of political influence, transparency and openness in a rather skeptical way: they usually described the political environment as having scant transparency, the availability of open data as low; they regarded the economic environment as unable to encourage forms of data journalism, and the Italian newsroom as poor in resources. In addition to these general features of the Italian media system, data journalists also define the general journalistic culture in Italy as opposed to their own values, which is often due to the fact that the interviewed data journalists have been trained and outside Italy by mainly Anglo-American educators, socialized in a completely different journalism culture.

Discussion and conclusion

Although data journalism is still a relatively minor phenomenon in Italy, several data journalism awards have been assigned to Italian projects, demonstrating that the field offers high quality work. For instance, Dataninja won a Data Journalism Award 2014 with the *The Migrants' Files* investigation. Yet, unlike in other European countries such as the UK, data journalism has not taken the entire journalistic field - or at least the main established news outlets - by storm.
We can, however, observe specific Italian trends: The negative economic and political contexts have strengthened the community and fostered a strong data journalists network. In addition, the absence of programmatic economic investments by traditional publishers force data journalists to excel, as freelancers, in a highly competitive market that aims 1) to publish in established Italian news media; 2) to publish in international news outlets, often in collaboration with international journalists; 3) to gain prizes and grants (e.g. from institutions like journalismgrant.org, journalismfund.eu, pewresearch.org) where high professional standards are required. The sociological approach adopted in this article has aimed precisely to analyze the enactment of the Italian forms of data journalism in relation to social institutions that lie outside journalism itself (Anderson 2013: 1009). What this article suggests is that Italian data journalism has grown *vis-à-vis* a disadvantageous political environment and outside the well-reputed news organizations. From this perspective, Italian data journalism is an anomaly precisely because it offers overall what can be considered as a good product, in a highly disadvantageous environment. In other words, the obstacles at the institutional and organizational levels have been a propellant to innovation, or at least to act beyond (sometimes against) certain constraints. The same discourse may be applied to the individual level: data journalism's lack of a *systematic* education by journalism schools, the *Ordine dei giornalisti* or higher education requires Italian data journalists to become highly professionalized *outside* the institutionalized journalistic field. In other words: one cannot enter the field if one has not acquired the tools, values and instruments to propose high quality products *on one's own*. In fact, only high-quality products can obtain awards and grants in an increasingly competitive context. Because data journalists in Italy are mainly freelancers, this seems to be the only way that data journalists can have their work acknowledged and appreciated by established news outlets, which are generally slow in adopting innovation. As professional data journalists are slowly but steadily entering the newsrooms of legacy news media as well, they may become change agents on their own by instituting new practices, workflows and norms.
The heterodox educational path, together with the need for specific skills in a highly competitive and professional sub-field may explain why, compared to the traditional

representation of Italian journalistic culture, data journalists adopt more rigorous professional standards particularly with regard to transparency. According to Reese and Shoemaker (2016), the boundaries of the media have shifted due to global connectivity and citizen interaction. This implies a more networked quality than the older conceptual partitioning of media and audience. In the case of Italian data journalism, this networked quality means also the enactment at the practice as well as at the routine level of the norm of transparency. We relate this distinct awareness of transparency specifically to the lack of a traditional and institutionalized path of data journalism education. People interested in becoming data journalists are forced to acquire new knowledge and expertise mostly on their own, also because many of them have a traditional journalistic background and have thus never encountered such new forms of journalism. By doing so, they will become aware of values, ideas and approaches mainly originating from an Anglo-American journalistic culture: not surprisingly most of the data journalists interviewed stressed their faith in the norm of objectivity, a principle widely accepted in Anglo-American journalism, but not so much in the Mediterranean journalism culture.

Furthermore, because many data journalism-related courses are accessible online, data journalists will acquire knowledge about new tools, practices and ideas influenced by digital culture, specifically by movements such as *networked journalism*, *open source*, *computing open culture* - generally, by a culture that places great value on concepts such as openness, sharing, collaboration and non-discrimination. With regard to the first research question, we can thus affirm that Italian data journalists often articulate data journalism as a practice in opposition to traditional journalism strongly influenced by politics, as being influenced by structural and institutional constraints such as a low level of innovation or a limited transparency of the government and the public administration and, at the same time, at the individual level of journalists, as a practice with a strong fact-orientation based on well-established principles such as objectivity and openness. The findings in Italy thus confirm those studies that have identified a strong influence of the open data movement on the development of data journalism in many other European countries (Porlezza 2019). In addition, the findings also reflect some of the previously carried out academic work with regard to collaborations across boundaries between data journalists and other social actors such as NGOs, data activists, computational experts in the community, and sometimes even whistleblowers in the public administration. Particularly with regard to the collection of specific data, where the resources in terms of time and finances exceed the possibilities of the analyzed agencies, forms of complementary practices of facilitating along the lines of Baack (2018) could be observed: while NGOs provide relevant data for the journalists in order to hold the government to account for their actions, the NGOs benefit from a the journalists' work of facilitating engagement and generating attention – ultimately a win-win-situation for both parties involved. However, also in the case if Italy, an increased openness of journalism does not come with a wider participation in the journalistic production. Italy is far from examples such as *La Nacion* in Argentina, which is one of the brightest in terms of citizen participation (see Palomo et al. 2019) In other words: albeit different forms of collaboration between data journalists and other social actors, the gatekeeping function remains firmly in the hands of the journalists. This is also reflected in the data journalists' understanding of openness: it is about allowing users to understand how journalistic labor works and how they can make sense of data. It is not about them participating in the news production process.

In light of our findings we can state that most Italian data journalists are practicing what they preach, particularly in terms of being transparent about their work. Particularly during the reconstruction interviews, journalists showed us in which part of their work they explained the sources that they used or the kind of data analysis that they implemented. Moreover, many of those journalists used *medium.com* to further explain their work to interested users - or their

peers - or they commented on their own work in the various social media groups in which they were active and involved. Even if transparency is still a challenge for data journalists in Italy – traditional publishers, for instance, still do not adapt to these norms due to the different cultural context in journalism, or because of the fear of free-riding by the competitors – they show a deep concern for being open and for providing the most reliable information in each step of the production process.

However, while data journalists show a keen commitment to transparency that traditional Italian newsrooms generally still struggle to apply, journalism's performance in terms of its ability to provide a certain benefit, i.e. transparency, allowing for criticism and control and conveying political knowledge, seems to be largely determined by the context in which it emerges. Particularly with regard to data journalism, one structural factor that exerts a strong influence on its performance is the availability and accessibility of public datasets. When considering the performance of Italian data journalists, it seems clear that their aim remains tied to the quality and openness of the data. Not surprisingly, those data journalists interviewed before the new FOIA was implemented, stated that their impact on the public administration in terms of increasing its transparency was close to zero. With regard to the second research question we can thus affirm that, within journalism, Italian data journalists implement transparency measures, but they are not able to perform similarly on the societal level. However, this constraint might not only be due to the limited transparency of the public administration, but also to the fact that – at least at the time when we carried out the study – not many legacy news media outlets employed data journalists, which limited the scope, the impact and the audience of data journalism.

Data journalism on its own, therefore, is not powerful enough to counteract the opaqueness of the public administration, even if there are leaks from time to time. But these leaks are not a consequence of data journalism performance. Nevertheless, journalists can somehow counteract the limited access to public datasets either by obtaining information through third parties such as NGOs, or they can collect data themselves, without relying on data from the public administration. However, given that the resources of freelance data journalists working for specialized agencies are limited, specific collaborations with NGOs, where the division of labor between data collection and analysis/publication are complementary, occur regularly. In any case, since May 2016 Italian journalists have also the right to submit FOIA requests.

Yet, even if FOIA-driven data have apparently increased the opportunities to hold the government and the public administration to account, the question whether data journalism is actually able to improve the openness of the public administration as well as the free flow of information in Italy also depends on whether the public administration understands that openness and transparency have positive side-effects, particularly in terms of citizen participation, as stated in the Open Government Declaration.

Obviously, this study has some limitations: first and foremost, the sampling method. However, because this an exploratory study on a relatively new phenomenon in Italy, the method chosen to identify experts was appropriate despite its clear shortcomings. In addition, the analysis presented can only provide an intermediate result, also because it was carried out in a very specific period when the Italian parliament was actually deliberating the first Italian FOIA. Furthermore, the data journalists in the sample were mostly working for independent and specialized agencies. In the meantime, as the community of data journalists is slowly growing, there may be professional data journalists working for established newsrooms as well, and they may have different perspectives on the topic.

However, in spite of these limitations, the significance of the findings should not be underestimated: the results still demonstrate that both the structural context and the educational paths within a given media system strongly influence data journalism's performance. Furthermore, it would also be useful to carry out comparative research into data journalism,

something that is still missing in this field of research. It would also be helpful to look more thoroughly into the specific ethical challenges of data journalism and determine, not only what values data journalists are actually upholding (as Italian data journalists are striving for objectivity) but also whether the current codes of ethics have to be rethought in terms of the use of (sensitive) data and the use of algorithms or even artificial intelligence when it comes to the processing of data within newsrooms.

Notes

[1] Among many indexes that assess the level of a country's digitization and evaluate both its level of access and e-government/open data facilities, see the Digital Economy and Society Index (DESI) issued by the European Commission, which summarizes relevant indicators on Europe's digital performance (Italy is 25th in the overall ranking); and the Open Data Barometer to estimate the level of openness of data provided by the public administration (Italian performances are better that the average of the rest of the countries, but Italy is far from UK or US, which, not surprisingly, present the best data journalism products).

[2] The direct quotes were translated into English by the authors. Each journalist is coded with a specific number. The low number of interviewees, the fact that most agencies are small sized, and the fact that we guaranteed anonymity prevent us from providing further details about the journalists.

[3] The fact that data journalists are closely interconnected, particularly on Twitter, was recently demonstrated by Splendore (2016b).

[4] The projects are: "Vaccini, la mappa di chi continua a rifiutarli. La situazione migliora, ma non molto", "L'Unione dei diritti", and "Quanti non pagano l'assicurazione dell'auto?" These stories are about the use of vaccines by Italians, the civic rights in the European Union and the payment of car insurance. They are available here:
https://www.wired.it/scienza/medicina/2017/11/07/vaccini-mappa-italia-rifiuta,
http://lab.gruppoespresso.it/espresso/2017/europa-unione-diritti-60-anni-trattati-roma and
http://www.infodata.ilsole24ore.com/2017/05/07/quanti-non-pagano-lassicurazione-dellauto-scopri-la-mappa-dei-comuni-la-media-italiana-12 (last access 25 January 2019).

[5] See the Open Government Declaration at https://www.opengovpartnership.org/open-government-declaration (last access 25 February 2019)

References

Agostini, Angelo. 2004. *Giornalismi. Media e giornalisti in Italia*. Bologna: Il Mulino.

Aitamurto, Tanja., Esa Sirkkunen, and Pauliina Lehtonen. 2011. *Trends in data journalism*. Espoo: VTT.

Ananny, Mike, and Kate Crawford. 2015. "A Liminal Press." *Digital Journalism* 3 (2): 192-208.

Anderson, Chris W. 2013. "Towards a Sociology of Computational and Algorithmic Journalism". *New Media & Society* 15 (7): 1005–1021.

Anderson, Chris W. and Juliette De Maeyer. 2015. "Objects of journalism and the news". *Journalism* 16 (1): 3–9.

Anderson, Chris. W. 2015. "Between the Unique and the Pattern: Historical Tensions in Our Understanding of Quantitative Journalism". *Digital Journalism* 3 (3): 349-363.

Appelgren, Ester, and Gunnar Nygren. 2014. "Data Journalism in Sweden: Introducing New Methods and Genres of Journalism into "Old" Organizations". *Digital Journalism* 2 (3): 394–405. 10.1080/21670811.2014.884344.

Appelgren, Ester, and Ramón Salaverría. 2018. "The Promise of the Transparency Culture: A Comparative Study of Access to Public Data in Spanish and Swedish Newsrooms". *Journalism Practice* 12 (8): 986–96. 10.1080/17512786.2018.1511823

Ausserhofer, Julian, Robert Gutounig, Michael Oppermann, Sarah Matiasek, and Eva Goldgruber. 2017. "The Datafication of Data Journalism Scholarship: Focal Points,

Methods, and Research Propositions for the Investigation of Data-Intensive Newswork". *Journalism*, online first 10.1177/1464884917700667
Baack, Stefan. 2015. "Datafication and empowerment: How the open data movement re-articulates notions of democracy, participation, and journalism. " *Big Data & Society* 2 (2): online only. 10.1177/2053951715594634
Baack, Stefan. 2018b. "Practically Engaged: The Entanglements Between Data Journalism and Civic Tech." Digital Journalism 6 (6): 673–692.
Becker, Howard. S. 1963. *Outsiders: Studies in the Sociology of Deviance*. New York: Macmillan.
Belair-Gagnon, Valerie, and Avery E. Holton. 2018. "Boundary Work, Interloper Media, and Analytics in Newsrooms: An Analysis of the Roles of Web Analytics Companies in News Production." Digital Journalism 6 (4): 492–508.
Bentivegna, Sara, and Rita Marchetti. 2018. "Journalists at a Crossroads: Are Traditional Norms and Practices Challenged by Twitter?" *Journalism* 19 (2): 270–90. 10.1177/1464884917716594.
Borges-Rey, Eddy. 2016. "Unravelling Data Journalism: A Study of Data Journalism Practice in British Newsrooms". *Journalism Practice* 10 (7): 833–43.
Borges-Rey, Eddy. 2017. "Towards an Epistemology of Data Journalism in the Devolved Nations of the United Kingdom: Changes and Continuities in Materiality, Performativity and Reflexivity". *Journalism*, online first. 10.1177/1464884917693864.
Boyles, Jan Lauren and Eric Meyer. 2016. "Letting the Data Speak, Digital Journalism", *Digital Journalism* 4 (7): 944-954.
Brugger, Jerome, Marianne Fraefel, Reinhard Riedl, Hansjakob Fehr, Daniel Schoeneck, and Christoph Staehli Weissbrod. 2016. "Current Barriers to Open Government Data Use and Visualization by Political Intermediaries." Conference proceedings of the 2016 Conference for E-Democracy and Open Government (CeDEM). 10.1109/CeDEM.2016.18
Cheruiyot, David, Stefan Baack, and Raul Ferrer-Conill. 2019. 'Data Journalism Beyond Legacy Media: The Case of African and European Civic Technology Organizations'. Digital Journalism, April, 1–15. https://doi.org/10.1080/21670811.2019.1591166.
Cheruiyot, David, and Raul Ferrer-Conill. 2018. "'Fact-checking Africa': Epistemologies, data and the expansion of journalistic discourse." *Digital Journalism,* online first. DOI: 10.1080/21670811.2018.1493940
Coddington, Mark. 2019. "Defining and mapping data journalism and computational journalism: A review of typologies and themes". In S. Eldridge II & B. Franklin (Eds.), *Routledge handbook of developments in digital journalism studies*. London: Routledge.
Coddington, Mark. 2014. "Clarifying journalism's quantitative turn: A typology for evaluating data journalism, computational journalism, and computer-assisted reporting". *Digital Journalism* 3 (3): 331-348.
Cornia, Alessio. 2014. "Will Italian political journalism ever change?" In R. Kuhn, and R. Kleis Nielsen (eds.) *Political Journalism in Transition: Western Europe in a Comparative Perspective*, 47–73. London: Tauris.
D'Arma, Alessandro. 2015. *Media and Politics in Contemporary Italy: From Berlusconi to Grillo.* Lexington Books: London.
De la Serna, Carlos Martínez. 2018. Collaboration and the creation of a new journalism commons. Available at: https://www.cjr.org/tow_center_reports/collaboration-and-the-journalism-commons.php
De Maeyer, Juliette, Manon Libert, David Domingo, François Heinderyckx, and Florence Le Cam. 2015. "Waiting for data journalism: A qualitative assessment of the anecdotal take-up of data journalism in French-speaking Belgium". *Digital Journalism* 3 (3): 432–46. 10.1080/21670811.2014.976415

Deuze, Mark, and Tamara Witschge. 2017. “Beyond journalism: Theorizing the transformation of journalism.” *Journalism* 19 (2): 165-181.
Edwards, Rosalind, and Janet Holland. 2013. *What Is qualitative interviewing?* London, New Delhi: Bloomsbury.
Eldridge II, Scott. A. 2018. *Online journalism from the periphery. Interloper media and the journalistic field.* New York: Routledge.
Ekström, Mats, and Oscar Westlund. 2019. “The Dislocation of News Journalism: A Conceptual Framework for the Study of Epistemologies of Digital Journalism”. Media and Communication 7 (1): 259. https://doi.org/10.17645/mac.v7i1.1763.
Felle, Tom. 2015. “Digital watchdogs? Data reporting and the news media’s traditional ‘fourth estate’ function”. *Journalism* 17 (1): 85-96.
Ferrer-Conill, Raul. 2017. “Quantifying journalism? A study on the use of data and gamification to motivate journalists”. *Television & New Media* 18 (8): 706–20. 10.1177/1527476417697271.
Fink, Katherine, and C. W. Anderson. 2014. “Data journalism in the United States: Beyond the ‘usual suspects’.” *Journalism Studies* 16 (4): 467–481.
Flew, Terry, Anne Daniel, and Christine Spurgeon. 2010. “The Promise of Computational Journalism”. In *Media, Democracy and Change*: *Refereed Proceedings of the Australian and New Zealand Communications Association Annual Conference,* 1-19, edited by McCallum, Kerry Canberra: Australia and New Zealand Communication Association.
Godler, Yigal, and Zvi Reich. 2017. ”Journalistic Evidence: Cross-Verification as a Constituent of Mediated Knowledge”. *Journalism* 18 (5): 558–574.
Graves, Lucas, and Magda Konieczna. 2015. “Sharing the News: Journalistic Collaboration As a Field Repair”. *International Journal of Communication* 9: 1966–1984.
Hallin Daniel, and Paolo Mancini. 2004. *Comparing Media Systems: Three Models of Media and Politics*. Cambridge: Cambridge University Press.
Hanitzsch, Thomas et al. 2011. “Mapping journalism cultures across nations. A comparative study of 18 countries.” *Journalism Studies* 12 (3): 273-293.
Hanna, Paula. 2012. “Using Internet technologies (such as Skype) as a research media: a research note.” *Qualitative research* 12 (2), 239-242.
Karlsen, Joakim, and Eirik Stavelin. 2014. “Computational Journalism in Norwegian Newsrooms”. *Journalism Practice* 8 (1): 34–48. 10.1080/17512786.2013.813190
Karlsson, Michael. 2010. “Rituals of transparency. Evaluating online news outlets’ uses of transparency rituals in the United States, United Kingdom and Sweden.” *Journalism Studies* 11 (4), 535-545.
Kennedy, Helen, Thomas Poell, and Jose van Dijck. 2015. “Data and Agency”. *Big Data & Society* 2 (2): online only. https://doi.org/10.1177/2053951715621569.
Kovach, Bill and Tom Rosenstiel. 2014. *Elements of Journalism. What Newspeople Should Know and What the Public Should Expect*. New York: Three Rivers Press.
Lewis, Seth C. 2012. “The tension between professional control and open participation. Journalism and its boundaries”. *Information, Communication & Society* 15 (6): 836-866.
Lewis, Seth C., and Nikki Usher. 2013. “Open Source and Journalism: Toward New Frameworks for Imagining News Innovation”. *Media, Culture & Society* 35 (5): 602-619.
Lewis, Seth C., and Nikki Usher. 2014. “Code, collaboration, and the future of journalism. A Case Study of the Hacks/Hackers Global Network”. *Digital Journalism* 2 (3): 383-393.
Lewis, Norman P., and Eisa Al Nashmi. 2019. “Data Journalism in the Arab Region: Role Conflict Exposed”. Digital Journalism, May, 1–15. https://doi.org/10.1080/21670811.2019.1617041.

Loosen, Wiebke, Julius Reimer, and Fenja De Silva-Schmidt. 2017. "Data-driven reporting: An on-going (r)evolution? An analysis of projects nominated for the *Data Journalism Awards* 2013-2016". *Journalism*. Online first. 10.1177/1464884917735691

Lowrey, Wilson, and Jue Hou. 2018. "All Forest, No Trees? Data Journalism and the Construction of Abstract Categories." *Journalism*. Online first. 10.1177/1464884918767577

Mancini, Paolo. 2005. Political complexity and alternative models of journalism. The Italian case. In J. Curran and M.-J. Park (eds.) *De-Westernizing Media Studies,* 265-278. London: Routledge.

Palomo, Bella, Laura Teruel, and Elena Blanco-Castilla. 2019. 'Data Journalism Projects Based on User-Generated Content. How La Nacion Data Transforms Active Audience into Staff'. Digital Journalism, June, 1–19. https://doi.org/10.1080/21670811.2019.1626257.

Parasie, Sylvain, and Eric Dagiral. 2012. "Data-Driven Journalism and the Public Good: 'Computer-Assisted-Reporters' and 'Programmer-Journalists' in Chicago." *New Media & Society* 15 (6): 853–871.

Porlezza, Colin. 2016. Dall'open journalism all'open government. Trasparenza, partecipazione e l'impatto del data journalism. [From open journalism to open government. Transparency, participation and the impact of data journalism.] *Problemi dell'informazione* XL (1): 167-194.

Porlezza, Colin. 2017. Data journalism: il giornalismo oltre il confine delle redazioni. [Data journalism: Journalism beyond the boundaries of the newsroom.] In M. Antenore, and S. Splendore (eds.) *Data Journalism. Concetti, strumenti, esperienze per una nuova informazione,* 36-45. Milano: Mondadori.

Porlezza, Colin. 2018. Deconstructing data-driven journalism. Reflexivity between the datafied society and the datafication of news work. *Problemi dell'informazione* XLIII (3): 369-392.

Porlezza, Colin. 2019. Data journalism and the ethics of open source. Transparency and participation as a prerequisite for serving the public good. In A. Daly, S. K. Devitt, and M. Mann (eds.) *Good Data,* 189-201. Amsterdam: Institute of Network Cultures.

Reese, Stephen D., and Pamela J. Shoemaker. 2016. A Media Sociology for the Networked Public Sphere: The Hierarchy of Influences Model. Mass Communication and Society 19(4): 389–410. https://doi.org/10.1080/15205436.2016.1174268.

Reuters Institute. 2019. Digital News Report 2019. Oxford: Reuters Institute for the Study of Journalism.

Rogers, Simon. 2012. "*Open Data Journalism*" Accessed September 7 2018. https://www.theguardian.com/news/datablog/2012/sep/20/open-data-journalism

Sandoval-Martín, Maria Teresa, and Leonardo La-Rosa. 2018. "Big Data as a differentiating sociocultural element of data journalism: the perception of data journalists and experts." *REV - Communication & Society* 31 (4): 193-209.

Singer, Jane B. 2007. "Contested autonomy: professional and popular claims on journalistic norms." *Journalism Studies* 8 (1): 79–95.

Shoemaker, Pamela J., and Stephen D. Reese. 2014. *Mediating the Message in the 21st Century: A Media Sociology Perspective*. Third edition. New York: Routledge/Taylor & Francis Group.

Splendore, Sergio. 2016a. "Quantitatively Oriented Forms of Journalism and Their Epistemology." *Sociology Compass* 10 (5): 342-352

Splendore, Sergio. 2016b. "Closed data! Il giornalismo italiano alla ricerca di dati" [Closed data! Italian journalism is looking for data], Problemi dell'Informazione, XL (1): 194-215

Splendore, Sergio. 2017. *Giornalismo ibrido. Come cambia la cultura giornalistica italiana* [Hybrid journalism. How Italian journalism is changing], Roma, Carocci.

Splendore, Sergio, Philip Di Salvo, Tobias Eberwein, Harmen Groenhart, Michal Kus, and Colin Porlezza. 2016. “Educational Strategies in the Field of Data Journalism – A Comparative Study in Six European Countries.” *Journalism* 17(1): 138-152.

Stoneman, Jonathan. 2015. “Does Open Data Need Journalism?”. Accessed September 7 2018. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/Stoneman%20-%20Does%20Open%20Data%20need%20Journalism.pdf

Tabary, Constance, Anne-Marie Provost, and Alexandre Trottier. 2016. “Data Journalism’s Actors, Practices and Skills: A Case Study from Quebec.” *Journalism* 17 (1): 66–84.

Tandoc, Edson C., Jr., and Soo-Kwang Oh. 2017. “Small Departures, Big Continuities? Norms,Values, and Routines in The Guardian’s Big Data Journalism.” *Journalism Studies* 18 (8): 997–1015.

Usher, Nikki. 2016. *Interactive Journalism. Hackers, Data, and Code*. Urbana: University of Illinois Press.

Valencia, Juan-Carlos and Paula Restrep. 2019. Truly Smart Cities. Buen Conocer, Digital Activism and Urban Agroecology in Colombia. In A. Daly, S. K. Devitt, and M. Mann (eds.) *Good Data*, 317-329. Amsterdam: Institute of Network Cultures.

Weihnacht, Stefan, and Ralf Spiller. 2014. “Datenjournalismus in Deutschland. Eine explorative Untersuchung zu Rollenbildern von Datenjournalisten.” *Publizistik* 59 (4): 411-433.

Young, Mary Lynn, Alfred Hermida, and Johanna Fulda. 2018. “What Makes for Great Data Journalism?” *Journalism Practice* 12 (1): 115–135.

Zamith, Rodrigo. 2019. “Transparency, Interactivity, Diversity,and Information Provenance in Everyday Data Journalism.“ *Digital Journalism*, online first 10.1080/21670811.2018.1554409

Zhang, Shuling, and Jieyun Feng. 2018. “A Step Forward?: Exploring the Diffusion of Data Journalism as Journalistic Innovations in China”. *Journalism Studies*, online first. 10.1080/1461670X.2018.1513814