

# Transparency in informative content. An analysis of methods in Spanish data journalism (2019-2022)

Félix Arias Robles  
Cristian Ramón Marín-Sanchiz  
Andrea Abellán-Mancheño  
José Alberto García-Avilés  
Universidad Miguel Hernández de Elche  
farias@umh.es  
c.marin@umh.es  
abellanandrea8@gmail.com  
jose.garciaa@umh.es



Submission date: June 2022  
Accepted date: May 2023  
Published in: June 2023

**Recommended citation:** ARIAS ROBLES, F., MARÍN-SANCHIZ, C. R., ABELLÁN-MANCHEÑO, A. and GARCÍA-AVILÉS, J.A. (2023). “Transparency in informative content. An analysis of methods in Spanish data journalism (2019-2022)”. *Anàlisi: Quaderns de Comunicació i Cultura*, 68, 97-116. DOI: <<https://doi.org/10.5565/rev/analisi.3548>>

## Abstract

The crisis of confidence in journalism has increased hopes of internal transparency. The technical possibilities of data journalism and its background in social sciences support the values of openness and transparency. This study focuses on the daily journalism produced by the data units of the Spanish media (n=10) between January 2019 and April 2022. Through analysis of quantitative and qualitative content of a random sample of articles (n=62), this study explores the transparency of sources and the use of a methodology section. The results reveal an unequal and, on many occasions, infrequent use of elements such as an explanation of the research process or the possibility for data download. A greater use is observed in native online media and small and specialised projects, with certain exceptions, reflecting the reluctance of some media to implement the new “transparency rituals” and the greater importance of attitude and individual initiative rather than resources. Finally, an ambivalent influence of COVID-19 is detected in these practices, which suggest a critical view of data journalism and the evolution of its transparency.

**Keywords:** data journalism; transparency; methodologies; social sciences; open source; COVID-19

**Resum.** *Transparència en els continguts informatius. Una anàlisi dels mètodes en el periodisme de dades espanyol (2019-2022)*

---

La crisi de confiança en el periodisme ha augmentat les esperances dipositades en la transparència interna dels continguts. Les possibilitats tècniques del periodisme de dades i la seva vinculació amb les ciències socials suposen una oportunitat per posar en pràctica aquests valors d'obertura i traçabilitat. Aquesta recerca se centra en les peces que, en el dia a dia (no projectes especials), van produir els mitjans espanyols amb unitats de dades ( $n = 10$ ) entre gener de 2019 i abril de 2022. A través d'una anàlisi de contingut quantitatiu i qualitatiu d'una mostra aleatòria d'articles ( $n = 62$ ), aquest estudi explora la transparència de les fonts i l'ús de la secció específica de metodologia. Els resultats revelen un ús desigual i moltes vegades poc freqüent d'elements, com l'explicació del procés de recerca o la possibilitat de descàrrega de dades. S'observa un desenvolupament més gran en els mitjans nadius digitals i en els projectes petits i especialitzats, amb unes quantes excepcions. Es constata així la reticència d'alguns mitjans a implementar els nous «rituals de transparència» i la major importància de l'actitud i la iniciativa individual enfront dels recursos de les redaccions. Finalment, es detecta una influència ambivalent de l'COVID-19 en aquestes pràctiques, la qual cosa suggereix la necessitat de fomentar una visió crítica del periodisme de dades i l'evolució de la transparència interna del periodisme.

**Paraules clau:** periodisme de dades; transparència; metodologies; ciències socials; codi obert; COVID-19

**Resumen.** *Transparencia en los contenidos informativos. Un análisis de métodos en el periodismo de datos español (2019-2022)*

---

La crisis de confianza en el periodismo ha aumentado las esperanzas depositadas en la transparencia interna de los contenidos. Las posibilidades técnicas del periodismo de datos y su vinculación con las ciencias sociales suponen una oportunidad para poner en práctica estos valores de apertura y trazabilidad. Esta investigación se centra en las piezas que, en el día a día (no proyectos especiales), produjeron los medios españoles con unidades de datos ( $n = 10$ ) entre enero de 2019 y abril de 2022. A través de un análisis de contenido cuantitativo y cualitativo de una muestra aleatoria de artículos ( $n = 62$ ), este estudio explora la transparencia de las fuentes y el uso de la sección específica de metodología. Los resultados revelan un uso desigual y, en muchas ocasiones, poco frecuente de elementos, como la explicación del proceso de investigación o la posibilidad de descarga de datos. Se observa un mayor desarrollo en los medios nativos digitales y en los proyectos pequeños y especializados, con algunas excepciones. Se constata así la reticencia de algunos medios a implementar los nuevos «rituales de transparencia» y la mayor importancia de la actitud y la iniciativa individual frente a los recursos de las redacciones. Finalmente, se detecta una influencia ambivalente del COVID-19 en estas prácticas, lo que sugiere la necesidad de fomentar una visión crítica del periodismo de datos y la evolución de la transparencia interna del periodismo.

**Palabras clave:** periodismo de datos; transparencia; metodologías; ciencias sociales; código abierto; COVID-19

---

## 1. Introduction

Content transparency is considered one of the main solutions to such pressing problems in journalism as information overload or lack of trust (Chadha and Koliska, 2015: 215; Diakopoulos and Koliska, 2017: 5; Karlsson, 2010: 535;

Plaisance, 2007: 187). It would improve authority (Singer, 2007: 79), credibility (Hayes, Singer and Ceppos, 2007: 262), differentiation from non-professional content (Karlsson, 2011: 279) and added value for users (Gray, Bounegru and Chambers, 2012: 22). Therefore, some journalists and news organisations view content transparency as a way to make journalism “more systematic, accurate and trustworthy” (Zamith, 2019: 1).

This internal transparency of journalism has limitations and is not a “silver bullet” (Karlsson, Clerwall and Nord, 2014: 668), as there are problems with conceptualization, appropriation by journalists, and practical application. This is why it is even more important to study its use in daily data journalism.

Although transparency is far from new to journalism, it has found a favourable breeding ground in recent specialisations. It has been studied in computer-assisted reporting, where transparency has intentionally been kept low (Coddington, 2015: 10); in fact-checking, incorporating more advanced methodologies (Beiler, Irmer and Breda, 2020: 4; Díez-Garrido and Farpón, 2020: 398); and in computational journalism, where coding presents challenges such as translating technical terms into understandable language (Calvo-Rubio and Ufarte-Ruiz, 2021: 673; Díaz-Campo and Chaparro-Domínguez, 2020; Dörr, 2016: 11; Shin, 2021: 1063).

But above all, transparency has been studied in relation to data journalism, a rare growth area in an industry battered by economic challenges (Zamith, 2019: 1). In Spain, its development began in 2013 (Antón-Bravo, 2013: 113; Chaparro-Domínguez, 2014; Ferreras-Rodríguez, 2013), coinciding with the creation of the transparency law (Appelgren and Salaverría, 2018: 9; La-Rosa and Sandoval-Martín, 2016: 1210; Tejedor Fuentes, 2014: 620). Three years later, Ferreras-Rodríguez (2016: 270) identified 14 Spanish media outlets that regularly or sporadically produce data journalism. In 2019, the specialisation was consolidated in legacy media with the creation of a data unit at *El País* (Rojas-Torrijos and García-Cepero, 2020).

The monitoring and analysis of COVID-19 gave a boost to this specialisation in Spain, as in the rest of the world (Córdoba-Cabús, García-Borrego and López-Martín, 2020: 333). The complexity of the situation demanded resources and time when news publishers were struggling to survive amid significant losses in advertising revenues that were not offset by increases in reader revenues (Westlund and Hermida, 2020). But this progress was not without certain limitations, especially due to dependence on official data (Wu, 2021) and lack of transparency (Díez-Garrido and Farpón, 2020: 397). This situation increased the need for a more critical reflection on the specialisation (Gray and Bounegru, 2021).

The consolidation of data journalism has not gone unnoticed in academia (Appelgren, Lindén, and Van-Dalen, 2019; Arias-Robles and López-López, 2020; Ausserhofer et al., 2017). “An explosion in data journalism-oriented scholarship” was noted some years ago (Fink and Anderson, 2015: 467), and soon included content transparency (Zamith, 2019: 2). However, data journalism is still evolving and remains a “largely unexplored” field in journalism

research, aside from national case studies, epistemological insights and ethnographies based on interviews with data journalists (Stalph, 2018: 1).

Therefore, it seems appropriate to look more closely into what data journalistic outputs look like (Loosen, Reimer and De Silva-Schmidt, 2017: 6). There is some research on data journalism content, which addresses transparency, including in the Spanish media context (Chaparro-Domínguez and Díaz-Campo, 2021: 11; Córdoba-Cabús, Huber and Farias-Batlle, 2023; Córdoba-Cabús and García-Borrego, 2021: 365), but it is often focused on large, internationally prize-winning projects. Chadha and Koliska (2015: 215) and Zamith (2019: 1) observed that comparatively few empirical studies have examined transparency in the context of daily news production.

## 2. Literature review

### 2.1. *Journalism and social sciences*

Inspiration from scientific methods was probably the origin of journalism's interest in transparency. At the beginning of the 20th century, Max Weber in Germany and Walter Lippmann and Robert E. Park in the United States already proposed the first synergies between the work of the journalist and that of the social scientist to overcome a crisis in the profession generated by the abuse of opinion (Dader, 1997). Most of these postulates crystallised with the implementation of the so-called precision journalism (PJ) (Meyer, 1973). Precision journalism entered newsrooms when journalists began to transform reports produced by experts such as sociologists into news, but especially when they experimented with methods from social and behavioural research (Cortés del Alamo, Elías and Luengo Cruz, 2018: 70).

Precision journalism soon evolved into computer-assisted reporting (CAR), data-driven journalism (DDJ) or computational journalism (CJ). All of these concepts incorporated new achievements such as open source or computer programming, but the basis continued to be in the inclusion of social sciences in the exercise of the journalism (Coddington, 2015; Sandoval-Martín and La-Rosa, 2018). This allowed journalists to go beyond the more conventional and impressionistic work based mainly on the juxtaposition of statements (Tong and Zuo, 2019).

The most common techniques come from disciplines such as sociology or anthropology, but also from other fields of knowledge, such as mathematics (López-García, Toural-Bran and Rodríguez-Vázquez, 2016). A recent study focused on the application of the scientific method in the data journalism project *Medicamentalía*, promoted by the specialised, digital-only outlet *Civio*, reveals various similarities with a scientific article, such as the presence of a methodology or the analysis of the results, which enhance transparency and information accuracy (Marín-Sanchiz and Arias-Robles, 2021).

Some authors doubt that data journalism work can always fulfil such standards. Weinacht and Spiller (2014: 412) argue that, unlike scientific

work, the data journalistic inquiry process often starts from the availability of data instead of from research questions. The selection of angles for reporting is influenced by news values and the public's interest in the respective field. Most of the literature highlights the danger of taking this relationship for granted, linking data journalism with greater objectivity (Beiler et al., 2020: 4; Tandoc and Oh, 2015: 997). Data journalism, like any specialisation, would only represent a part of reality, although sometimes a larger one, and it does not avoid possible biases (Anderson, 2018).

## *2.2. Transparency in journalism*

Although not new, transparency has become an important value in today's society, from government and businesses to journalism and computer sciences. The metaphor of transparency encompasses three virtues to identify the motives behind actions and to improve trust: informational substantiality, accountability and participation (Balkin, 1999). The first two are closely related to the focus of this research and to the recently highlighted seven steps of transparency in statistical practice, which include assessing data pre-processing choices and sharing data and code (Wagenmakers et al., 2021).

This is accentuated in the current digital context, where open-source and hacker culture assume a "transparent rubric" in which source code is shared (Lewis and Usher, 2013: 608). And it has been assumed by journalism, often codified in professional codes of ethics (Córdoba-Cabús and García-Borrego, 2021), even as an alternative to the classic normative demands of objectivity, increasingly questioned since the 1980s (Vos and Moore, 2020: 18).

Kovach and Rosenstiel (2007: 83) argued for journalism to be as transparent as possible about its methods and motives. Explaining how the story came to be and why it was presented the way it was would be critical to the development of a more discerning public. Deuze (2005: 456) argues for transparency in journalism to give everybody a chance to monitor, check, criticise and even become involved in the journalistic process. Moreover, Plaisance (2007: 187) defined transparency as an ethical imperative and as much of a mainstay of journalism as credibility, honesty and fairness.

There were some limits to keep in mind. Lewis and Usher detailed off-the-record confidentiality or state secrecy, and highlighted that open-source journalism does not mean revealing and documenting every conversation, but invites the reader to better understand the process: "Radical transparency is neither called for nor necessarily desirable" (2013: 614). According to this idealistic view, since media are facing increased scrutiny of their products, the best way to respond would be letting people see the process that leads to the creation of those products (Allen, 2008: 324).

Transparency was technically achievable but mostly absent in the analogue media system, as this was shaped by a closed news culture (Deuze, 2003: 539). However, the actual application of these postulates in practice was more limited than in theory (Lowrey and Hou, 2018: 8; Zhang and Feng, 2019: 14). Trans-

lating transparency into a commonly enacted ritual was not common (Singer, 2007). Sometimes, it was due to a conscious action of journalists. Chadha and Koliska (2015) revealed that, while most of the discussions around transparency in the news media are normative, many journalists have not yet embraced transparency. Hellmueller, Vos and Poepsel (2013: 288) add that some journalists still prefer objectivity as a guiding journalistic principle. According to Plaisance and Skewes (2003: 842), journalists often prioritise other values over transparency. Singer (2007) argued that journalists may see transparency as an intrusion on their autonomy, and Karlsson (2010) found that transparency is often routinized into a strategic ritual that separates execution from intent by promoting a small degree of transparency, but resisting too much of it.

These results contrast with one of the few surveys on transparency carried out among journalists. Beiler et al. (2020: 13) found that two thirds of the respondents from daily newspapers (66%) and seven out of ten from public broadcasters agree “rather strongly” that data journalism makes journalism more transparent. The practitioners of data journalism believe in their craft’s ability to transform journalism towards more transparency. Similarly, journalists interviewed by Ferreras-Rodríguez (2016: 261-262) support the convenience of publishing raw data and explaining the methodologies used to promote transparency. This highlights the divergence between what they think should happen and what they actually do. As we will discuss below, most recent research underscores the poor application of transparency in journalistic content. Only Felle (2016: 2) found that transparency is reflected in actions such as the publication of content methodologies that evidence “the strengthened accountability role” of this specialisation.

This is why many researchers are calling for a rethink of the concept, and view transparency as a normative goal, a consolidated routine, rather than an instrumental value. Calling for more transparency without consideration of what is trying to be accomplished might present more problems than solutions. Thus, journalists should not adopt transparency as a way of improving public standing or increasing legitimacy, mainly because it probably will not work (Allen, 2008: 333-336). To be useful, Karlsson (2010: 535-536) argued that journalists should adopt “rituals of transparency” that can be communicated, understood and accepted by audiences and colleagues.

Framed in this context, the main objective of this paper is to analyse the evolution and peculiarities of transparency in the content produced by the data journalism units of the Spanish media, through the following research questions:

- RQ1. What transparency-related features are used in Spanish data journalism pieces in the period from January 2019 to April 2022?
- RQ2. Are there differences in transparency levels between traditional Spanish news outlets and digital-native projects?
- RQ3. Has the COVID-19 pandemic increased transparency levels in Spanish data journalism?

### 3. Methodology

This paper uses a mixed method approach, similar to previous research (Beiler et al., 2020; Loosen et al., 2017; Stalph, 2018). First, we did a preliminary study of Spanish news outlets that have a data unit, through a directory and survey of more than a hundred journalists gathered from a previous research project (Arias and Carvajal, 2022). Forty-eight journalists from ten media outlets were identified.

**Table 1.** Analysed media outlets

Media outlet	Root	Focus	Period
Civio	Digital	Data journalism	2019 (S1) - 2022 (S1)
Datadista	Digital	Data journalism	2020 (S1) - 2022 (S1)
El Confidencial	Digital	Generalist	2019 (S1) - 2022 (S1)
eldiario	Digital	Generalist	2019 (S1) - 2022 (S1)
El Mundo	Print	Generalist	2019 (S1) - 2022 (S1)
El País	Print	Generalist	2019 (S1) - 2022 (S1)
La Sexta	Television	Generalist	2020 (S1) - 2022 (S1)
Maldita	Digital native	Fact checking	2019 (S1) - 2022 (S1)
Newtral	Digital native	Fact checking	2019 (S1) - 2022 (S1)
RTVE	Television	Generalist	2021 (S1) - 2022 (S1)

Source: Authors' own.

Next, all the articles published by these journalists between January 2019 and April 2022 were recorded from the author's pages in their media. When these compilations did not exist (RTVE) or there were no by-lines on the contents (Maldita), the media archive and their Twitter accounts – an important platform for data journalists, used by Felle (2016: 5) and Zhang (2018: 737) – were used.

From this collection, a random selection was applied. There were only two conditions: avoiding repetition in the same semesters, and including pieces that were part of the news organisations' own research (no replication of data). The aim was to obtain content from before, during, and after the COVID-19 pandemic, a key moment for the evolution of data journalism (Desai et al., 2021).

Finally, one article from each of the ten media outlets in the seven semesters was analysed, except the eight in which there is no data. Because some media did not publish data journalism articles in 2019 and 2020, a final sample of 62 articles was obtained. The sample was subjected to quantitative and qualitative analysis. Using a Google Spreadsheet as a database, two different coders went over every article to characterise it, based on the following variables:

- Data sources. We detailed if the article appeared with the specific name of the source and if it is linked at the text and graphic footnotes.

This is what Leon (2021: 81) considers elementary even in works based on outside data.

- Methodology. After assessing whether or not there was a specific section, usually at the end, with the details of the journalistic work, two features were registered: a logbook or description of the details of the process followed by journalists, with special attention to the relationship with the sources and the limitations of the result; and a description of the variables or concepts used. Mazotte (2021: 87) points out that this is a practice already consolidated in most data-driven research, and is appropriate when journalists account for their own journalistic production process.
- Downloadable data. We differentiated between partial data, belonging to a concrete graphic, and the complete datasets. Including the complete database is the best way to guarantee the reproducibility of the research (Henkel et al., 2020: 1). In the most advanced phase, the inclusion of the programming code used (Leon, 2021: 81) was ruled out because it was not present in any of the studies analysed.

#### 4. Results

One of the cornerstones of transparency in journalism is citing sources. Results show that virtually all the pieces analysed (98%) explain the origin of the different information. However, these sources were not always accessible, as only three news outlets (Civio, Maldita, Newtral and RTVE) provided links to the sources in each text analysed. The media with the greatest room for improvement were El Mundo and El País.

**Table 2.** Sources included in text

Media outlet	2019		2020				2021				2022				TOTAL	
	S1		S2		S1		S2		S1		S2		S1		F	L
	N	L	N	L	N	L	N	L	N	L	N	L	N	L		
Civio															100%	100%
Datadista															100%	80%
El Confidencial															100%	86%
El Mundo															100%	29%
El País															86%	57%
eldiario															100%	71%
La Sexta															100%	60%
Maldita															100%	100%
Newtral															100%	100%
RTVE															100%	100%
<b>TOTAL</b>	100%	86%	86%	86%	100%	78%	100%	67%	100%	70%	100%	80%	100%	80%	98%	77%

(Key: blue = sources included; white = not included; yellow = no data. "S" = semester; "N" = name of the source; and "L" = the inclusion of a link)

Source: Authors' own.



Sources are specified less frequently in graphics (Table 3). 79% of the pieces in the sample specify a source, but only six outlets include sources in all graphics, while 21% of them provide a link. For instance, Datadista and eldiario included zero links in their graphics, while Maldita (29%), Newtral (29%) incorporated them only in some of their articles.

**Table 3.** Sources included in graphics

Media outlet	2019		2020				2021				2022				TOTAL	
	S1		S2		S1		S2		S1		S2		S1		F	L
	N	L	N	L	N	L	N	L	N	L	N	L	N	L		
Civio															43%	14%
Datadista															100%	0%
El Confidencial															100%	43%
El Mundo															43%	14%
El País															43%	0%
eldiario															100%	0%
La Sexta															100%	40%
Maldita															100%	29%
Newtral															86%	29%
RTVE															100%	67%
<b>TOTAL</b>	<b>71%</b>	<b>43%</b>	<b>57%</b>	<b>0%</b>	<b>78%</b>	<b>0%</b>	<b>89%</b>	<b>22%</b>	<b>80%</b>	<b>40%</b>	<b>80%</b>	<b>20%</b>	<b>90%</b>	<b>20%</b>	<b>79%</b>	<b>21%</b>

(Key: blue = sources included; white = not included; yellow = no data. "S" refers to the semester)

Source: Authors' own.

Only 24% of the content analysed includes a methodology section. In fact, there are cases (La Sexta, El Mundo, Maldita, Newtral and RTVE) that do not explain their methods in a clearly bounded piece of content. The media with highest rates are Civio (86%) and El País (57%), and generally, the methods section specify the sources the content is built up on (93,3%).

**Table 4.** Existence of a methodology

Media outlet	2019		2020		2021		2022		TOTAL
	S1	S2	S1	S2	S1	S2	S1		
Civio									86%
Datadista									20%
El Confidencial									14%
El Mundo									0%
El País									57%
eldiario									14%
La Sexta									0%
Maldita									29%
Newtral									0%
RTVE									0%
<b>TOTAL</b>	<b>29%</b>	<b>29%</b>	<b>11%</b>	<b>33%</b>	<b>20%</b>	<b>30%</b>	<b>20%</b>	<b>20%</b>	<b>24%</b>

(Key: blue = methodology included; white = not included; yellow = no data. "S" refers to the semester)

Source: Authors' own.

Methodological sections usually consist of plain text with a few links to sources or e-mail addresses. They are often placed at the end of the content. Moreover, they are commonly written in the second person plural.

**Figure 1.** Excerpt from a methodology published in El País

### **Metodología**

A continuación detallamos las fuentes y explicamos el proceso seguido para localizar cada servicio en el mapa. Es importante tener en cuenta que el análisis es una aproximación: las listas no son completamente exhaustivas y habrá errores en algunos municipios.

**Hospitales.** La lista de hospitales en el mapa proviene de [Catálogo Nacional de Hospitales](#) de 2018, elaborado por el Ministerio de Sanidad. Contiene más de 900 hospitales en toda España, con su dirección completa. Para ubicarlos en el mapa hemos geolocalizado cada hospital con un proceso automático que se alimenta de Google Maps.

Source: elpais.com.

These methodology sections are also characterised by a wide diversity in length. Although the general average is 331 words, Civio averaged 533 words over five pieces, while El País used an average of 168 to explain their methods. As a curiosity, El País published an article whose methodology was longer than the content itself.<sup>1</sup> The longest methodology sections were in Civio (1222 words) and El Mundo (1007), and the shortest in Datadista (48) and El País (44).

**Table 5.** Number of words in the methodology (“S” means the Semester)

Media outlet	2019		2020		2021		2022	TOTAL	AVERAGE
	S1	S2	S1	S2	S1	S2	S1		
Civio	236		360	599	1222		250	2667	533
Datadista			106	48				154	77
El Confidencial		110						110	110
El Mundo	1007							1007	1007
El País		373		44		147	107	671	168
eldiario	111							111	111
La Sexta									
Maldita	725				66	149		940	313
Newtral									
RTVE									
<b>TOTAL</b>	2079	483	466	691	1288	296	357	809	331
<b>AVERAGE</b>	520	242	233	230	644	148	179		

Source: Authors' own.

1. <[https://elpais.com/politica/2019/10/25/actualidad/1572027354\\_718725.html](https://elpais.com/politica/2019/10/25/actualidad/1572027354_718725.html)>.

In most cases the methodology section lacks a “logbook” that would allow the reader to understand the steps followed by the data journalists. In fact, only 23% of the descriptions analysed include this item, and five organisations (El Mundo, eldiario, La Sexta, Newtral and RTVE) do not provide them in any content. Interestingly, Datadista, a firm specialising in data-journalism, shows a 20% rate.

**Table 6.** Description of the process included in the methodology

Media outlet	2019		2020		2021		2022	TOTAL
	S1	S2	S1	S2	S1	S2	S1	
Civio	Blue	White	Blue	White	Blue	White	Blue	86%
Datadista	Yellow	Yellow	Blue	White	White	White	White	20%
El Confidencial	White	Blue	White	White	White	White	White	14%
El Mundo	White	White	White	White	White	White	White	0%
El País	White	White	White	Blue	White	White	Blue	43%
eldiario	White	White	White	White	White	White	White	0%
La Sexta	Yellow	Yellow	White	White	White	White	White	0%
Maldita	White	White	Blue	White	Blue	White	Blue	43%
Newtral	White	White	White	White	White	White	White	0%
RTVE	Yellow	Yellow	Yellow	Yellow	Yellow	Yellow	Yellow	0%
<b>TOTAL</b>	14%	14%	33%	22%	20%	30%	20%	23%

(Key: blue = description included; white = not included; yellow = no data. “S” refers to the semester)

Source: Authors' own.

Explanations of variables (type of data, if the databases were built ad hoc for the articles, etc.) show a similar result. Datadista, El Mundo, La Sexta, Newtral and RTVE did not detail them, while Maldita (14%), eldiario and El Confidencial (14%) did so in a few cases. El País (43%) and Civio (86%) display higher rates of description of variables.

**Table 7.** Explanation of variables

Media outlet	2019		2020		2021		2022	TOTAL
	S1	S2	S1	S2	S1	S2	S1	
Civio	Blue	White	Blue	White	Blue	White	Blue	86%
Datadista	Yellow	Yellow	White	White	White	White	White	0%
El Confidencial	White	Blue	White	White	White	White	White	14%
El Mundo	White	White	White	White	White	White	White	0%
El País	White	White	White	White	White	White	Blue	43%
eldiario	Blue	White	White	White	White	White	White	14%
La Sexta	Yellow	Yellow	White	White	White	White	White	0%
Maldita	White	White	White	White	White	Blue	White	14%
Newtral	White	White	White	White	White	White	White	0%
RTVE	Yellow	Yellow	Yellow	Yellow	Yellow	Yellow	Yellow	0%
<b>TOTAL</b>	29%	29%	11%	11%	10%	30%	20%	19%

(Key: blue = explanation included; white = not included; yellow = no data. “S” refers to the semester)

Source: Authors' own.

Some articles also offer the possibility to download partial data (26%) or even the whole dataset (11%). El Confidencial included the partial data feature in all the articles in the sample. Datadista (40%), La Sexta (40%) and Newtral (43%) show similar rates. El Mundo (29%) also featured this functionality, but it was available on Civio, El País, eldiario, Maldita or RTVE. Again, news outlets rely on the technological possibilities offered by online services like Datawrapper or Flourish.

**Table 8.** Partial and full data download

Media outlet	2019		2020		2021		2022		TOTAL				
	S1		S2		S1		S2		S1				
	P	T	P	T	P	T	P	T	P	T			
Civio											0%	86%	
Datadista											40%	0%	
El Confidencial											100%	0%	
El Mundo											29%	0%	
El País											0%	0%	
eldiario											0%	0%	
La Sexta											40%	0%	
Maldita											0%	14%	
Newtral											43%	0%	
RTVE											0%	0%	
<b>TOTAL</b>	14%	14%	29%	0%	44%	11%	22%	11%	30%	10%	10%	25%	10%

(Key: blue = download included; white = not included; yellow = no data. "S" refers to the semester)  
Source: Authors' own.

Civio can be considered to be a benchmark of the willingness to offer a complete dataset, as they link to a new page in the methodology. Here, they publish the source, time range and format, as well as a thorough explanation of the dataset. To download the dataset, the user's name, email address and organisation are requested, and they are asked to cite and link to Civio.

**Figure 2.** Civio's dataset download page

**Pruebas diagnósticas para detectar casos de coronavirus**  
Actualización con el número (total y relativo) de pruebas realizadas por comunidades autónomas para diagnosticar casos de covid-19

---

Fuente **Ministerio de Sanidad y comunidades autónomas**

Rango temporal **Febrero 2020 - abril 2021**

Formato **XLSX**

Laboratorios de toda España realizan a diario miles de pruebas para diagnosticar nuevos casos de infección por **coronavirus**. La falta de suficientes análisis para detectar personas afectadas por la **covid-19** fue uno de los grandes problemas para atrapar a tiempo al virus SARS-CoV-2.

Durante semanas, el Gobierno rechazó dar datos actualizados sobre el número de pruebas diagnósticas realizadas. Solo desde el 23 de abril de 2020, el Ministerio de Sanidad publica las PCR, los test de antígenos y las pruebas serológicas de forma semanal. Ante esta falta de transparencia, Civio decidió ir actualizando a diario los

Si eres socio o socia de Civio y quieres que te ayudemos a entender, tratar o buscar datos relevantes, escríbenos al [buzón de correo](#). Estaremos encantados de echarte una mano.

**descarga estos datos**

Reutiliza nuestros datos de forma libre y gratuita. [Las reglas son claras](#): cita y enlaza. Y si encuentras algún error, por favor, avísanos.

Nombre

Email

Source: Civio.es.

Last, we observed the presence of a specific mailbox dedicated to communicating with users who want to provide new information or help amend errors. This is the least common feature, appearing in only 10% of the pieces. Maldita is the leader in this regard (43%), followed by El Confidencial (29%) and El País (14%). The offerings of Civio, Datadista, El Mundo, eldiario, LaSexta, Newtral and RTVE lack this feature.

## 5. Discussion

This study assessed transparency in daily content produced by Spanish data journalism teams from 2019 to 2022. Specifically, it sheds light on which transparency-related features are used (RQ1): they always indicate the source of their data, but do not link to them. Major news producers in the Spanish media ecosystem, such as El Mundo and El País, were among those that provided fewer links. A similar problem can be observed in the inclusion of sources in graphics and figures (e.g., maps).

Only 24% of the content analysed includes a methodology section, which is a hallmark of data journalism. Content often lacked an explanation of variables or a process outline to describe the steps that journalists had followed in obtaining and analysing datasets. Although it is clear that this practice is common of data journalism (Felle, 2016), the results are consistent with previous research on the topic: Díez-Garrido and Farpón (2020: 414) revealed that fewer than two of the twelve Latin American media outlets they analysed openly published databases. Chaparro-Domínguez and Díaz-Campo (2021: 11) also found that award-winning data journalism projects complied with verification and data analysis, but transparency and privacy principles were followed to a lesser extent.

The results also reveal low rates of downloadable data, which echoes Zamith's (2019) findings that only 13% of stories offered data downloads. The general absence of this added value reflects that, beyond the additional costs that transparency involves (Granados and Gupta, 2013: 640), the shortcomings in practice are mainly explained by the prioritisation of other issues (Plaisance and Skewes, 2003: 842) and by the fear of an intrusion into journalistic work (Singer, 2007). And this confirms how far transparency has to go until it is routinely adopted by journalists (Singer, 2007) or becomes a regular routine (Chadha and Koliska, 2015). Karlson's (2010) views of resistance of many journalists to incorporate high levels of openness in their content were confirmed.

Of course, there are nuances to these deficits and in many cases they depend on the type of media outlet (RQ2). In general, digitally-native companies are more likely to introduce transparency initiatives, and they lead in including methodologies, explanations of variables, the possibility to download data, and linking sources in graphics. This trend is particularly pronounced among small, specialised journalism initiatives such as Civio. However, not all digital-first organisations are more transparent than traditional

media companies, as the case of El País, a large legacy media outlet, shows. These exceptions reflect the complexity of organisational culture in newsrooms, and the different speed at which innovations are adopted (Porcu, Hermans and Broersma, 2020). This forces to refine the scant open culture attributed to analogue media (Deuze, 2003: 539). It also reinforces the idea of data units as autonomous entities in large newsrooms, and data journalists as agents of innovation, sometimes without the support of their company or environment (Appelgren et al., 2019: 1191; Arias and Carvajal, 2022: 21; Ausserhofer et al., 2017: 17-18; Baack, 2018: 674; Cheruiyot, Baack and Ferrer-Conill, 2019: 13; Mutsvauro, 2019: 1289; Wright and Nolan, 2021).

The COVID-19 pandemic, although considered an important period for data journalism, had a limited effect on transparency practices (RQ3). Our research shows that the description of journalistic processes was lower in 2019, and that the possibility of downloading data reached its peak in 2020, and maintained high figures over the last two years. However, the results on the percentage of links to sources or the existence of methodology sections do not seem to have increased as a result of the health crisis. The length of these appendices had the lowest number of words in 2020, in theory the hardest year in terms of deaths, hospitalizations and restrictions. This lack of growth, despite the rise of data journalism during the pandemic, is probably explained by the dependence on official data (Wu, 2021). Díez-Garrido and Farpón (2020: 397) already highlighted the lack of transparency despite the abundance of data in this period. The precarious situation of journalists due to the lack of resources and insecure working conditions under which many of these contents were produced also played an important role.

This study has some limitations. First, the inclusion of cases from a single country, although with important peculiarities: Spain was the last European country to pass a transparency law, and a dozen data journalism units have been established in media organisations over the last ten years. Second, the sample size is very limited and not representative; a large-scale mapping of Spanish data journalism output may offer more nuances on the different transparency initiatives.

However, this is an exploratory, pilot study that for the first time looks closely at the Spanish market holistically. It includes a qualitative perspective in the detailed analysis of specific practices, but semi-structured interviews would have been advisable, to obtain a more thorough understanding of the data journalism units' processes and decisions.

To overcome these limitations, we encourage future researchers to conduct semi-structured interviews with data journalists and editors, to better understand their strategies; and also to perform a cross-national study to compare the work of more news outlets and understand the influence of the media ecosystems in relation to transparency. Additionally, it is proposed that future content analysis research also looks at the identification and linking of sources; the specific elements of the methodological section; and the downloading of complete data. But analysis is also recommended regarding the

presence of the relevant programming code, which, as indicated above, has been omitted because it does not appear in any of the pieces analysed. This element will have more value in the future use of artificial intelligence tools.

In conclusion, transparency initiatives in Spanish data journalism must be boosted, although the simple application of a few parameters is not a sufficient condition to increase actual transparency. For this reason, this study supports the idea of creating “useful, everyday rituals of transparency” (Karls-son, 2010: 535-536) with the aim of defining an authentic system of transparency, accountability and legitimacy of news media. Integrating these practices into the professional culture of journalists, such as including references to sources, will be key to fostering the credibility and sustainability of the media. Showing “the back room” of information production in the methodology and disseminating the complete data of the investigation will differentiate quality journalism from false or biased content.

To achieve this, a training-focused approach must be considered, both in the heart of news outlets and in university programs (Lewis, McAdams and Stalph, 2020; Burns and Matthews, 2018). As Gray and Bounegru (2021) point out, this reality should lead us to be more reflective and critical in analysing the quality of data journalism development.

### Bibliographical references

- ALLEN, D. S. (2008). “The trouble with transparency: The challenge of doing journalism ethics in a surveillance society”. *Journalism Studies*, 9 (3), 323-340.  
<<https://doi.org/10.1080/14616700801997224>>
- ANDERSON, C. W. (2018). *Apostles of Certainty: Data Journalism and the Politics of Doubt*. Oxford: Oxford University Press.
- ANTÓN-BRAVO, A. (2013). “El periodismo de datos y la web semántica”. *CIC Cuadernos de Información y Comunicación*, 18, 99-116.  
<[https://doi.org/10.5209/rev\\_CIYC.2013.v18.41718](https://doi.org/10.5209/rev_CIYC.2013.v18.41718)>
- APPELGREN, E., LINDÉN, C. G. and VAN DALEN, A. (2019). “Data Journalism Research: Studying a Maturing Field across Journalistic Cultures, Media Markets and Political Environments”. *Digital Journalism*, 7 (9), 1191-1199.  
<<https://doi.org/10.1080/21670811.2019.1685899>>
- APPELGREN, E. and SALAVERRÍA, R. (2018). “The Promise of the Transparency Culture: A comparative study of access to public data in Spanish and Swedish newsrooms”. *Journalism Practice*, 12 (8), 986-996.  
<<https://doi.org/10.1080/17512786.2018.1511823>>
- ARIAS-ROBLES, F. and CARVAJAL-PRIETO, M. (2022). “Periodistas de datos. Análisis del perfil sociodemográfico del profesional en España y Latinoamérica”. *Estudios Sobre El Mensaje Periodístico*, 28 (1), 13-24.  
<<https://doi.org/10.5209/esmp.76395>>
- ARIAS-ROBLES, F. and LÓPEZ-LÓPEZ, P. J. (2020). “Driving the Closest Information. Local Data Journalism in the UK”. *Journalism Practice*, 1-13.  
<<https://doi.org/10.1080/17512786.2020.1749109>>
- AUSSERHOFER, J., GUTOONIG, R., OPPERMAN, M., MATIASEK, S. and GOLDGRUBER, E. (2017). “The datafication of data journalism scholarship: Focal points,

- methods, and research propositions for the investigation of data-intensive newswork". *Journalism*, 21 (7), 950-973.  
<<https://doi.org/10.1177/1464884917700667>>
- BAACK, S. (2018). "Practically Engaged. The entanglements between data journalism and civic tech". *Digital Journalism*, 6 (6), 676-692.  
<<https://doi.org/10.1080/21670811.2017.1375382>>
- BALKIN, J. M. (1999). "How mass media simulate political transparency". *Cultural Values*, 3 (4), 393-413.  
<<https://doi.org/10.1080/14797589909367175>>
- BEILER, M., IRMER, F. and BRENDA, A. (2020). "Data Journalism at German Newspapers and Public Broadcasters: A Quantitative Survey of Structures, Contents and Perceptions". *Journalism Studies*, 21 (11), 1571-1589.  
<<https://doi.org/10.1080/1461670X.2020.1772855>>
- CALVO-RUBIO, L. M. and UGARTE-RUIZ, M. J. (2021). "Artificial intelligence and journalism: Systematic review of scientific production in Web of Science and Scopus (2008-2019)". *Communication and Society*, 34 (2), 159-176.  
<<https://doi.org/10.15581/003.34.2.159-176>>
- CHADHA, K. and KOLISKA, M. (2015). "Newsrooms and Transparency in the Digital Age". *Journalism Practice*, 9 (2), 215-229.  
<<https://doi.org/10.1080/17512786.2014.924737>>
- CHAPARRO-DOMÍNGUEZ, M. Á. (2014). "Nuevas formas informativas: el periodismo de datos y su enseñanza en el contexto universitario/New information methods: data journalism and its teaching in the university context". *Historia y Comunicación Social*, 19, 43-54.  
<[https://doi.org/10.5209/rev\\_HICS.2014.v19.45009](https://doi.org/10.5209/rev_HICS.2014.v19.45009)>
- CHAPARRO-DOMÍNGUEZ, M. Á. and DÍAZ-CAMPO, J. (2021). "Data Journalism and Ethics: Best Practices in the Winning Projects (DJA, OJA and Sigma Awards)". *Journalism Practice*.  
<<https://doi.org/10.1080/17512786.2021.1981773>>
- CHERUIYOT, D., BAACK, S. and FERRER-CONILL, R. (2019). "Data Journalism Beyond Legacy Media: The case of African and European Civic Technology Organizations". *Digital Journalism*, 7 (9), 1215-1229.  
<<https://doi.org/10.1080/21670811.2019.1591166>>
- CODDINGTON, M. (2015). "Clarifying Journalism's Quantitative Turn: A typology for evaluating data journalism, computational journalism, and computer-assisted reporting". *Digital Journalism*, 3 (3), 331-348.  
<<https://doi.org/10.1080/21670811.2014.976400>>
- CÓRDOBA-CABÚS, A. and GARCÍA-BORREGO, M. (2021). "Evaluating the transparency in reference data journalism. Study of the stories published between 2018 and 2019". *Icono14*, 19 (2), 364-387.  
<<https://doi.org/10.7195/RI14.V19I2.1648>>
- CÓRDOBA-CABÚS, A., GARCÍA-BORREGO, M. and LÓPEZ-MARTÍN, Á. (2020). "El periodismo de datos durante la crisis sanitaria del COVID-19 en la prensa española". *Revista Ibérica de Sistemas y Tecnologías de Información*, 35, 325-337.
- CÓRDOBA-CABÚS, A., HUBER, B. and FARIAS-BATLLE, P. (2023). "Data journalism in Spain and Austria: features, organizational structure, limitations, and future perspectives". *El Profesional de La Información*, 32 (1).  
<<https://doi.org/10.3145/epi.2023.ene.16>>



- CORTÉS DEL ALAMO, H. M., ELÍAS, C. and LUENGO CRUZ, M. (2018). "Periodismo de datos y transparencia al margen de los grandes medios". *Revista Icono14*, 16 (2), 66-87.  
<<https://doi.org/10.7195/ri14.v16i2.1177>>
- DADER, J. L. (1997). *Periodismo de Precisión: La vía socioinformática de descubrir noticias*. Madrid: Síntesis.
- DESAI, A., NOUVELLET, P., BHATIA, S., CORI, A. and LASSMANN, B. (2021). "Data journalism and the COVID-19 pandemic: opportunities and challenges". *The Lancet. Digital Health*, 3 (10), e619-e621.  
<[https://doi.org/10.1016/S2589-7500\(21\)00178-3](https://doi.org/10.1016/S2589-7500(21)00178-3)>
- DEUZE, M. (2003). "The web and its journalisms: Considering the consequences of different types of newsmedia online". *New Media and Society*, 5 (2), 203-230.  
<<https://doi.org/10.1177/1461444803005002004>>
- (2005). "What is journalism? Professional identity and ideology of journalists reconsidered". *Journalism*, 6 (4), 442-464.  
<<https://doi.org/10.1177/1464884905056815>>
- DIAKOPOULOS, N. and KOLISKA, M. (2017). "Algorithmic Transparency in the News Media". *Digital Journalism*, 5 (7), 809-828.  
<<https://doi.org/10.1080/21670811.2016.1208053>>
- DÍAZ-CAMPO, J. and CHAPARRO-DOMÍNGUEZ, M.-Á. (2020). "Periodismo computacional y ética". *Revista Icono14*, 18 (1), 10-32.  
<<https://doi.org/10.7195/ri14.v18i1.1488>>
- DÍEZ-GARRIDO, M. and FARPÓN, C. R. (2020). "The institutional and media transparency of the coronavirus. An analysis of data portals and digital media in Ibero-America". *Revista Latina de Comunicación Social*, 78, 393-418.  
<<https://doi.org/10.4185/RLCS-2020-1482>>
- DÖRR, K. N. (2016). "Mapping the field of Algorithmic Journalism". *Digital Journalism*, 4 (6), 700-722.  
<<https://doi.org/10.1080/21670811.2015.1096748>>
- FELLE, T. (2016). "Digital watchdogs? Data reporting and the news media's traditional 'fourth estate' function". *Journalism*, 17 (1), 85-96.  
<<https://doi.org/10.1177/1464884915593246>>
- FERRERAS-RODRÍGUEZ, E. M. (2013). "Aproximación teórica al perfil profesional del 'Periodista de datos'". *Icono14*, 11 (2), 115.  
<<https://doi.org/10.7195/ri14.v11i2.573>>
- (2016). "El periodismo de datos en España". *Estudios sobre el Mensaje Periodístico*, 22 (1), 255-272.  
<[https://doi.org/10.5209/rev\\_ESMP.2016.v22.n1.52594](https://doi.org/10.5209/rev_ESMP.2016.v22.n1.52594)>
- FINK, K. and ANDERSON, C. W. (2015). "Data Journalism in the United States: Beyond the 'usual suspects'". *Journalism Studies*, 16 (4), 467-481.  
<<https://doi.org/10.1080/1461670X.2014.939852>>
- GRANADOS, N. and GUPTA, A. (2013). "Transparency Strategy: Competing with Information in a Digital World". *MIS Quarterly: Management Information Systems*, 37 (2), 637-641.
- GRAY, J., BOUNEGRU, L. and CHAMBERS, L. (2012). *The Data Journalism Handbook 1*. Sebastopol: O'Reilly Media. Retrieved from: <[https://datajournalismhandbook.org/uploads/first\\_book/DataJournalismHandbook-2012.pdf](https://datajournalismhandbook.org/uploads/first_book/DataJournalismHandbook-2012.pdf)>.
- GRAY, J. and BOUNEGRU, L. (2021). *The Data Journalism Handbook 2*. European Journalism Centre. Retrieved from: <<https://datajournalism.com/read/handbook/two>>.

- HAYES, A. S., SINGER, J. B. and CEPOS, J. (2007). "Shifting Roles, Enduring Values: The Credible Journalist in a Digital Age". *Journal of Mass Media Ethics*, 22 (4), 262-279.  
<<https://doi.org/10.1080/08900520701583545>>
- HELLMUELLER, L., VOS, T. P. and POEPEL, M. A. (2013). "Shifting Journalistic Capital?: Transparency and objectivity in the twenty-first century". *Journalism Studies*, 14 (3), 287-304.  
<<https://doi.org/10.1080/1461670X.2012.697686>>
- HENKEL, I., THURMAN, N., MÖLLER, J. and TRILLING, D. (2020). "Do Online, Offline, and Multiplatform Journalists Differ in their Professional Principles and Practices? Findings from a Multinational Study". *Journalism Studies*, 21 (10), 1-21.  
<<https://doi.org/10.1080/1461670X.2020.1749111>>
- KARLSSON, M. (2010). "Rituals of transparency: Evaluating online news outlets' uses of transparency rituals in the United States, United Kingdom and Sweden". *Journalism Studies*, 11 (4), 535-545.  
<<https://doi.org/10.1080/14616701003638400>>
- (2011). "The immediacy of online news, the visibility of journalistic processes and a restructuring of journalistic authority". *Journalism*, 12 (3), 279-295.  
<<https://doi.org/10.1177/1464884910388223>>
- KARLSSON, M., CLERWALL, C. and NORD, L. (2014). "You ain't seen nothing yet transparency's (Lack of) effect on source and message credibility". *Journalism Studies*, 15 (5), 668-678.  
<<https://doi.org/10.1080/1461670X.2014.886837>>
- KOVACH, B. and ROSENSTIEL, T. (2007). *The Elements of Journalism: What Newspeople Should Know and the Public Should Expect*. New York: Three Rivers Press.
- LA-ROSA, L. and SANDOVAL-MARTÍN, T. (2016). "La insuficiencia de la Ley de Transparencia para el ejercicio del Periodismo de datos en España". *Revista Latina de Comunicación Social*, 71, 1208-1230.  
<<https://doi.org/10.4185/RLCS-2016-1142>>
- LEON, S. (2021). "Accounting for Methods in Data Journalism: Spreadsheets, Scripts and Programming Notebooks". In: GRAY, J. and BOUNEGRU, L. (Eds). *The Data Journalism Handbook 2*. European Journalism Centre, 80-86. Retrieved from <<https://datajournalism.com/read/handbook/two>>
- LEWIS, S. and USHER, N. (2013). "Open source and journalism: Toward new frameworks for imagining news innovation". *Media, Culture and Society*, 35 (5), 602-619.  
<<https://doi.org/10.1177/0163443713485494>>
- LEWIS, N. P., MCADAMS, M. and STALPH, F. (2020). "Data Journalism". *Journalism and Mass Communication Educator*, 75 (1), 16-21.  
<<https://doi.org/10.1177/1077695820904971>>
- LOOSEN, W., REIMER, J. and DE SILVA-SCHMIDT, F. (2017). "Data-driven reporting: An on-going (r)evolution? An analysis of projects nominated for the Data Journalism Awards 2013-2016". *Journalism: Theory, Practice and Criticism*.  
<<https://doi.org/10.1177/1464884917735691>>
- LÓPEZ-GARCÍA, X., TOURAL-BRAN, C. and RODRÍGUEZ-VÁZQUEZ, A.-I. (2016). "Software, estadística y gestión de bases de datos en el perfil del periodista de datos". *El Profesional de La Información*, 25 (2), 286-294.  
<<https://doi.org/10.3145/epi.2016.mar.16>>

- LOWREY, W. and HOU, J. (2018). "All forest, no trees? Data journalism and the construction of abstract categories". *Journalism*.  
<<https://doi.org/10.1177/1464884918767577>>
- MARÍN-SANCHIZ, C. R. and ARIAS-ROBLES, F. (2021). "El periodismo (de datos) es de ciencias: estrategias para incrementar la calidad del contenido a través del método científico". In: DE LARA-GONZÁLEZ, A. and GÓMEZ MARTÍNEZ, Á (Eds.). *Ciencia y Periodismo. Una es de Marte y otra es de Venus*. Universidad Miguel Hernández de Elche. Retrieved from <<https://www.researchgate.net/publication/349862759>>.
- MAZOTTE, N. (2021). "Working Openly in Data Journalism". In: GRAY, J. and BOUNEGRU, L. (Eds.). *The Data Journalism Handbook 2*. European Journalism Centre, 87-89. Retrieved from <<https://datajournalism.com/read/handbook/two>>.
- MEYER, P. (1973). *Precision Journalism: A Reporter's Introduction to Social Science Methods* (Vol. 1). Lanham: Rowman & Littlefield Publishers.
- MUTSVAIRO, B. (2019). "Challenges Facing Development of Data Journalism in Non-Western Societies". *Digital Journalism*, 7 (9), 1289-1294.  
<<https://doi.org/10.1080/21670811.2019.1691927>>
- PLAISANCE, P. L. (2007). "Transparency: An Assessment of the Kantian Roots of a Key Element in Media Ethics Practice". *Journal of Mass Media Ethics*, 22 (2-3), 187-207.  
<<https://doi.org/10.1080/08900520701315855>>
- PLAISANCE, P. L. and SKEWES, E. A. (2003). "Personal and professional dimensions of news work: Exploring the link between journalist's values and roles". *Journalism and Mass Communication Quarterly*, 80 (4), 833-848.  
<<https://doi.org/10.1177/107769900308000406>>
- PORCU, O., HERMANS, L. and BROERSMA, M. (2020). "Unlocking the Newsroom: Measuring Journalists' Perceptions of Innovative Learning Culture". *Journalism Studies*, 21 (10), 1420-1438.  
<<https://doi.org/10.1080/1461670X.2020.1758956>>
- ROJAS-TORRIJOS, J. L. and GARCÍA-CEPERO, J. (2020). "Percepción del periodismo deportivo de datos entre usuarios habituales. Estudio de caso del modelo predictivo de El País para el Mundial de Fútbol de 2018". *Revista Mediterránea de Comunicación*, 11 (2), 1-16.  
<<https://doi.org/10.14198/MEDCOM2020.11.2.8>>
- SANDOVAL-MARTÍN, T. and LA-ROSA, L. (2018). "El big data como elemento sociocultural diferenciador del periodismo de datos". *Communication & Society Journal*, 31 (4), 193-209.  
<<https://doi.org/10.15581/003.31.4.193-209>>
- SHIN, D. (2021). "A cross-national study on the perception of algorithm news in the east and the west". *Journal of Global Information Management*, 29 (2), 102-117.  
<<https://doi.org/10.4018/JGIM.2021030105>>
- SINGER, J. B. (2007). "Contested autonomy: Professional and popular claims on journalistic norms". *Journalism Studies*, 8 (1), 79-95.  
<<https://doi.org/10.1080/14616700601056866>>
- STALPH, F. (2018). "Classifying Data Journalism: A content analysis of daily data-driven stories". *Journalism Practice*, 12 (10), 1332-1350.  
<<https://doi.org/10.1080/17512786.2017.1386583>>

- TANDOC, E. C. and OH, S. K. (2015). "Small Departures, Big Continuities?: Norms, values, and routines in The Guardian's big data journalism". *Journalism Studies*, 18 (8), 997-1015.  
<<https://doi.org/10.1080/1461670X.2015.1104260>>
- TEJEDOR FUENTES, L. (2014). *La transparencia del patrimonio e intereses económicos de los miembros del ejecutivo y legislativo bajo el análisis del "periodismo de precisión y de datos"*. Doctoral Thesis, Universidad Complutense de Madrid.
- TONG, J. and ZUO, L. (2019). "The Inapplicability of Objectivity: Understanding the Work of Data Journalism". *Journalism Practice*, 1-17.  
<<https://doi.org/10.1080/17512786.2019.1698974>>
- VOS, T. P. and MOORE, J. (2020). "Building the journalistic paradigm: Beyond paradigm repair". *Journalism*, 21 (1), 17-33.  
<<https://doi.org/10.1177/1464884918767586>>
- WAGENMAKERS, E. J. et al. (2021). "Seven steps toward more transparency in statistical practice". *Nature Human Behaviour*, 5 (11), 1473-1480.  
<<https://doi.org/10.1038/s41562-021-01211-8>>
- WEINACHT, S. and SPILLER, R. (2014). "Datenjournalismus in Deutschland". *Publizistik*, 59 (4), 411-433.  
<<https://doi.org/10.1007/s11616-014-0213-5>>
- WESTLUND, O. and HERMIDA, A. (2020). "Data journalism and misinformation". In: H. TUMBER and S. WAISBORD (Eds.). *The Routledge Companion to Media Disinformation and Populism*. London: Routledge, 142-150.  
<<https://doi.org/10.4324/9781003004431-16>>
- WRIGHT, S. and NOLAN, D. (2021). "The Development of Data Journalism in China: Influences, Motivations and Practice". *Digital Journalism*, 1-18.  
<<https://doi.org/10.1080/21670811.2021.1927779>>
- WU, S. (2021). "Data Objectivity in a Time of Coronavirus: Uncovering the Potential Impact of State Influence on the Production of Data-Driven News". *Digital Journalism*, 9 (9), 1231-1248.  
<<https://doi.org/10.1080/21670811.2021.1942111>>
- ZAMITH, R. (2019). "Transparency, Interactivity, Diversity, and Information Provenance in Everyday Data Journalism". *Digital Journalism*, 7 (4), 470-489.  
<<https://doi.org/10.1080/21670811.2018.1554409>>
- ZHANG, X. (2018). "Visualization, Technologies, or the Public?: Exploring the articulation of data-driven journalism in the Twittersphere". *Digital Journalism*, 6 (6), 737-758.  
<<https://doi.org/10.1080/21670811.2017.1340094>>
- ZHANG, S. and FENG, J. (2019). "A Step Forward? Exploring the diffusion of data journalism as journalistic innovations in China". *Journalism Studies*, 20 (9), 1281-1300.  
<<https://doi.org/10.1080/1461670X.2018.1513814>>