Data journalism in Spain and Austria: features, organizational structure, limitations, and future perspectives

Article *in* El Profesional de la Informacion · January 2023

DOI: 10.3145/epi.2023.ene.16

CITATION
1

READS
71

3 authors, including:

Alba Córdoba-Cabús
University of Malaga
30 PUBLICATIONS **102** CITATIONS

SEE PROFILE

Data journalism in Spain and Austria: features, organizational structure, limitations, and future perspectives

Alba Córdoba-Cabús; Brigitte Huber; Pedro Farias-Batlle

Nota: Este artículo se puede leer en español en:
https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87127

Recommended citation:

Córdoba-Cabús, Alba; **Huber, Brigitte**; **Farias-Batlle, Pedro** (2023). "Data journalism in Spain and Austria: features, organizational structure, limitations, and future perspectives". *Profesional de la información*, v. 32, n. 1, e320116.
https://doi.org/10.3145/epi.2023.ene.16

Article received on September 09th 2022
Approved on November 11th 2022

Alba Córdoba-Cabús ✉
https://orcid.org/0000-0002-3519-0583

Universidad de Málaga
Departamento de Periodismo
León Tolstoi, s/n
29071 Málaga, Spain
albacordoba@uma.es

Brigitte Huber
https://orcid.org/0000-0002-9070-4962

IU International University of Applied Sciences, Munich, Germany
University of Vienna, Austria
brigitte.huber1@iu.org
brigitte.huber@univie.ac.at

Pedro Farias-Batlle
https://orcid.org/0000-0002-9634-5198

Universidad de Málaga
Departamento de Periodismo
León Tolstoi, s/n
29071 Málaga, Spain
farias@uma.es

Abstract

This paper makes an important contribution to comparative research by examining data journalism in Spain and Austria. This paper examines the practice of data journalism from a triple perspective: (a) the common features of day-to-day work, (b) the organizational structure and the role of the teams in newsrooms, and (c) the obstacles to and the future of data-driven reporting. Results from content analysis of data-driven news stories in *El país* and *Der standard* (*N* = 136) show differences and similarities in the covered topics, sources, narrative style, visualizations, interactive functions, and levels of transparency. Interestingly, only 36.8% of the analyzed news stories correspond to the normative expectations of transparency by incorporating both sources and methodological details. While the Spanish newspaper shows significantly higher levels of transparency compared with the Austrian newspaper, both newspapers perform very similarly when it comes to providing access to raw data, which was the case in only every fifth news story analyzed. Findings from focused interviews with the heads of data journalism teams deliver interesting insights into specific challenges that each news outlet is facing when creating day-to-day data-driven news stories. This research confirms the relevance that data journalism has achieved in countries such as Spain and demonstrates the effort of journalists in countries without access to information and transparency laws to create data-driven stories.

Keywords

Data journalism; Digital journalism; Sources; Data visualization; Transparency; Interactivity; Journalists; Journalistic profession; Digital media; Newspapers; *El país*; *Der standard*; Spain; Austria.

Funding

This research was supported by the *Spanish Government* through a research stay grant at the *University of Vienna*, Austria (FPU2017/02141). This research is a result of the project "Use of social networks for information purposes among young Spaniards: incidental consumption of news, technological constraints and credibility of journalistic content" (Ref. PID2019-106932RB-100), funded by the Spanish *Ministry of Science and Innovation*.

1. Introduction

Journalism has always relied on data to inform, report, combine sources, or contextualize. However, the connection of digitalization and datafication paved the way for a new era for information professionals (**Mayer-Schönberger**; **Cukier**, 2013). Those changes led to "journalism's quantitative turn" (**Coddington**, 2015), and it reinforced professional profiles such as the data journalist with an "increased focus on measurement, outcomes assessment" (**Anderson**, 2015, p. 363).

The first reference to the term "data journalism" was made more than 15 years ago, when **Holovaty** (2006) mentioned the need to work with spreadsheets to collate the information they gathered. This specialization illustrated how the media environment was changing and how data represented and influenced society (**Van-Dijck**, 2014; **Porlezza**, 2019). **Arrese** (2022) insists that the essence of data journalism comes from economic journalism, especially because of its purpose: to reduce uncertainty and guide decision-making. Despite identifying successive initiatives, the term began to become popular, especially from 2010 onward with *Wikileaks* and Tim Berners-Lee's statements, in which he assured that the future of journalism lay in data analysis (**Howard**, 2014; **Uskali**; **Kuuti**, 2015). The same year, *The economist* published a special issue called *Data, data everywhere*, in which it compared data with oil, as one of the most valuable resources (**Toonders**, 2014). At first glance, there are no significant differences between data journalism and other previous disciplines such as precision journalism and computer-assisted journalism. However, the main distinction lies in the big data context faced by the journalist and the needs arising from it (**Gray**; **Bounegru**; **Chambers**, 2012; **Sandoval-Martín**; **La-Rosa**, 2018; **De-Lima-Santos**; **Mesquita**, 2021).

Data-driven journalism practice and the corresponding academic literature have grown significantly since 2013 (**Ausserhofer** *et al*., 2017; **Zamith**, 2019; **Heravi**; **Lorenz**, 2020). The vast majority of papers are limited to specific geographical areas, thus offering a segmented view. Although there has been a major increase in academic attention on underrepresented regions (e.g., **Mustvairo**; **Bebawi**; **Borges-Rey**, 2020; **Pereira**; **Mastrella**, 2022; **Camaj**; **Martin**; **Lanosga**, 2022; **Kashyap**; **Mishra**; **Bhaskaran**, 2022), research has mainly focused on countries with a long tradition in data journalism (**Appelgren**; **Lindén**; **Van-Dalen**, 2020; **De-Lima-Santos**; **Schapals**; **Bruns**, 2020) such as the United Kingdom (e.g., **Knight**, 2015; **Borges-Rey**, 2016; **Arias-Robles**; **López-López**, 2020; **Tong**, 2021) and the United States (e.g., **Parasie**; **Darigal**, 2012; **Fink**; **Anderson**, 2015; **Zamith**, 2019). For this reason, **Fink** and **Anderson** (2015), **Splendore** *et al.* (2016), **Cheruiyot**, **Baack** and **Ferrer-Conill** (2019), and **Wright**, **Zamith**, and **Bebawi** (2019) advocate for cross-national comparative studies to provide in-depth understanding of the development of this specialization and to explore the nuances of the phenomenon.

A comparison between Spain and Austria seems particularly suitable for this purpose: both are democratic societies with well-developed media systems. Little specific research on Spanish data journalism has been published, and moreover especially focused on specific topics such as politics, without delving into the features of day-to-day data stories and professional routines (**Rojas-Torrijos**; **Rivera-Hernández**, 2016; **Teruel-Rodríguez**; **Blanco-Castilla**, 2016; **Córdoba-Cabús**, 2018; **Córdoba-Cabús**; **García-Borrego**; **López-Martín**, 2020; **Rubio-Jordán**, 2021; **Arias-Robles**; **Carvajal**, 2022). Likewise, there are no studies related to Austrian data journalism, so this paper aims to address this research need and provide findings to mitigate this gap.

According to the categorization established by **Hallin** and **Mancini** (2004), Spain belongs to the polarized pluralist model and Austria to the democratic corporatist model. Nevertheless, data show how these countries converge in terms of a high concentration of media owners, the decline of newspaper circulation, and the restrictions on access to data due to limitations from legislation (**Figl**, 2017; **Salaverría**; **Gómez-Baveiredo**, 2017; **Grünangerl**; **Trappel**; **Tomaz**, 2021). **Albarrán** (2009) and **Ohlsson** (2015) already introduced the idea of the hybridization of models and the European trend toward homogenization.

The core objective of this study is to explore data journalism practice in Austria and Spain. More specifically, this study aims to (1) assess the content of data journalism in Spain and Austria by analyzing data-driven news stories in the Spanish newspaper *El país* and the Austrian newspaper *Der standard* and (2) explore work routines as well as the obstacles to and future scenarios of data-driven reporting by interviewing the heads of data journalism of these news outlets.

1.1. Data journalism and access to information

Data journalism is a relatively young field, so achieving an exhaustive definition is rather challenging. In the literature, several attempts to define data journalism can be identified. **Howard** (2014, p. 4), for instance, considered data journalism to be a process associated with data science that encompasses "gathering, cleaning, organizing, analysing, visualizing, and publishing data to support the creation of acts of journalism." Meanwhile, **Veglis** and **Bratsas** (2017, p. 111) specified that data journalists engage in the process of

> "writing articles based on the information and embedding visualizations… in the articles that help readers to understand the significant of the story or allow them to pinpoint data that relate to them."

Heravi (2018, p. 3) stated that data journalism consists of

> "finding stories in data –stories that are of interest to the public– and presenting these stories in the most appropriate manner for public use and reuse."

Following Coddington's definition, herein, we understand data journalism to be the result of a convergence between different knowledge areas and practices, i.e., as a

> "hybrid form [of journalism] that encompasses statistical analysis, computer science, visualization and web design, and reporting" (**Coddington**, 2015, p. 335).

"Empirical studies have shown that data-driven news stories strongly rely on government and institutional sources"

Three common components can be discerned in all of these: the exploration of databases, the proccessing of data, and the communication of results.

Focusing on the features of data journalism projects, empirical studies have shown that data-driven news stories strongly rely on government and institutional sources (**Karlsson**, 2010; **Parasie**; **Darigal**, 2012; **Knight**, 2015; **De-Maeyer** *et al.*, 2015; **Cushion**; **Lewis**; **Callaghan**, 2017; **Loosen**; **Reimer**; **Schmidt**, 2017; **Tandoc**; **Oh**, 2017; **Lowrey**; **Hou**, 2018; **Stalph**, 2018; **Young**; **Hermida**; **Fulda**, 2018; **Appelgren**, 2018; **Hyder**; **Nahid**, 2019; **Zamith**, 2019; **Zhang**; **Feng**, 2019; **Córdoba-Cabús**, 2020; **Rubio-Jordán**, 2021). The main purpose of data-driven stories is to inform and present data through narration and visualization (**Loosen**; **Reimer**; **Schmidt**, 2017; **Ojo**; **Heravi**, 2018; **Young**; **Hermida**; **Fulda**, 2018; **Córdoba-Cabús**, 2020; **Rubio-Jordán**, 2021). The author-driven style is the most common narrative structure applied (**Loosen**; **Reimer**; **Schmidt**, 2017; **Ojo**; **Heravi**, 2018). Regarding transparency, prior research has found that data stories only rarely show transparency elements such as access to the raw data (**Lowrey**; **Hou**, 2018; **Young**; **Hermida**; **Fulda**, 2018; **Stalph**, 2018; **Zhang**; **Feng**, 2019; **Zamith**, 2019; **Porlezza**; **Splendore**, 2019; **Córdoba-Cabús**; **García-Borrego**, 2021; **Chaparro-Domínguez**; **Díaz-Campo**, 2021). Although the findings in terms of visualization in data-driven projects are mixed, static figures and limited use of interactive functions seem common (**Loosen**; **Reimer**; **Schmidt**, 2017; **Tandoc**; **Oh**, 2017; **Young**; **Hermida**; **Fulda**, 2018; **Appelgren**, 2018; **Stalph**, 2018; **Ojo**; **Heravi**, 2018; **Appelgren**, 2018; **Córdoba-Cabús**; **García-Borrego**, 2020).

The development of data journalism depends heavily on the context in which it emerges. The establishment of transparency and access-to-information laws encourages the emergence of data-driven projects (**Parasie**; **Darigal**, 2012; **Appelgren**; **Nygren**, 2014; **Uskali**; **Kuutti**, 2015; **De-Maeyer** *et al*., 2015; **Tabary**; **Provost**; **Trottier**, 2016; **Appelgren**; **Salaverría**, 2018; **Sandoval-Martín**; **La-Rosa**, 2018; **Porlezza**; **Splendore**, 2019; **Tong**; **Zuo**, 2021; **Camaj**; **Martin**; **Lanosga**, 2022; **Kashyap**; **Mishra**; **Bhaskaran**, 2022; **Tong**, 2022). Hence, the most well-developed and successful data journalism is found in countries with laws that ensure compliance with these principles (**Howard**, 2014; **Cortés-del-Álamo**; **Luengo**; **Elías**, 2018).

In the European Union, regulations regarding access to public documents were adopted in 2001, but the implementation of these regulations differs by country (**Appelgren**; **Salaverría**, 2018). Thereafter, the *European Commission* adopted the public sector information directive in 2003 (*Directive 2003/98/EC*), which encouraged the reuse of public data and enabled access in digital format.

In Spain, the law on transparency, access to public information, and good governance was approved in 2013, and work is currently underway to draft the corresponding regulation together with civil society. However, this law is not in line with *European Court of Human Right*'s jurisprudence, and it lacks control mechanisms to ensure compliance (**Sánchez-Calero**; **Mancinas-Chávez**, 2017; **Appelgren**; **Salaverría**, 2018). In Austria, the constitution guarantees freedom of information, but there is no specific law regulating transparency. *The European Convention on Human Rights* guarantees access to information, but the Austrian constitution limits this right by requiring that institutions apply discretion (**Figl**, 2017; **Grünangerl**; **Trappel**; **Tomaz**, 2021). This may explain why Austria was at the bottom of the global right to information ranking in 2022, with 33 points out of 150 (*RTI*, 2022). In this regard, it is considered to be one of the weakest countries in the world. Meanwhile, Spain scored 73 out of 150 points.

In Austria, data journalists "are still a rare species" (**Figl**, 2017, p. 5). Until a few years ago, data journalism was only practiced in small newsrooms such as *Dossier*, the public broadcaster *ORF*, and the national newspaper *Der standard*. Officially, in 2021, the *Austrian Press Agency* (*APA*) set up the "*Data + Graphics*" team, which brings together experts in data, infographics, design, and web developers. This is expected to provide visibility and value to the discipline. The first examples of data journalism in Spain emerged within organizations such as *Fundación Civio* or *Medialab-Prado* in 2011. *El confidencial* was the first media outlet to set up a data journalism team in 2013. However, the main generalist newspaper in Spain (*El país*) did not publish data journalism stories until 2017, while the public broadcaster (*RTVE*) set up their data team in 2019. Research focused on Spain shows a progressive upward trend in data journalism. However, except for digital-native media such as *El confidencial*, Spanish newspapers did not commit to data journalism beyond specific projects until 2015. The health pandemic associated with the coronavirus highlighted the relevance of data journalism, as well as the rise of stories based on figures in Spain (**Córdoba-Cabús**; **García-Borrego**; **López-Martín**, 2020; **Costa-Sánchez**; **López-García**, 2020; **Sanahuja**; **López**, 2021).

To explore data journalism content and practices in Spain and Austria, we formulated the following broad research questions:

"The establishment of transparency and access-to-information laws encourages the emergence of data-driven projects"

RQ1. What are the common features of data-driven news stories published in *El país* and *Der standard*?

RQ1a. What are the narrative features of day-to-day data journalism published in *El país* and *Der standard*?

RQ1b. What is the level of transparency of data journalism stories in *El país* and *Der standard*?

RQ1c. What kind of visualizations do data journalism articles in *El país* and *Der standard* include?

RQ2. How has data journalism developed in *El país* and *Der standard*?

RQ3. How do data journalists of *El país* and *Der standard* see the future of data journalism?

2. Methods

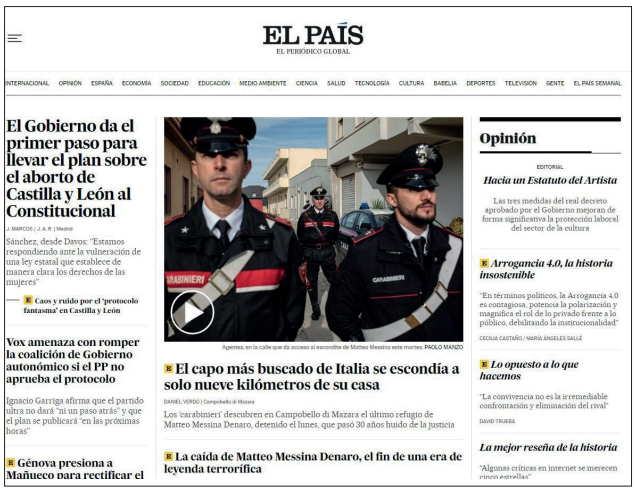
To answer these research questions, we combined quantitative and qualitative methods. First, we conducted a quantitative content analysis of data-driven news stories published in *El país* (Spain) and *Der standard* (Austria) to provide an overview of how data are used in news stories from these news outlets (*N* =136). Second, we employed interviews with two members of the data journalism teams of *El país* (Spain) and *Der standard* (Austria) to better understand the process of how journalists select and create news stories using data-driven techniques.

2.1. Content analysis

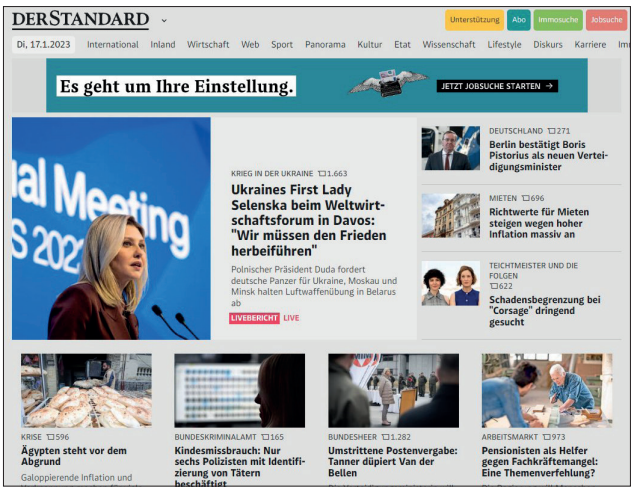
The quantitative content analysis focused on exploring the full content of data-driven news stories published online from 1 January to 30 June 2021, in the newspapers *El país* (Spain) and *Der standard* (Austria) (*N* = 136). Content analysis as a methodological tool enables one to understand the structure of stories as well as their basic components and operation (**Igartua**, 2006). These two newspapers were selected because they are the national newspapers most trusted by the population –48% and 69%, respectively– according to the *Digital news report* (*Reuters Institute*, 2021). They both have a data team in their newsroom and the highest online reach (21% and 20%, respectively) in the last year (*Reuters Institute*, 2021). We created a codebook consisting of formal categories (medium, date of publication, headline, section, and link) and categories that aimed to capture the content of data-driven news stories. Specifically, the proposed categories can be grouped into three dimensions: story features, data, and visualizations (Table 1).

Table 1. List of categories

Story features	
Topic	Politics; society; economy and business; health, science, and environment; education; sport; culture and art; entertainment; other
Number of personal sources	
Personal sources	No personal sources; government or public institutions; academics and experts; entrepreneurs; civil society; culture, art, and sports; ordinary people; other
Attribution	On the record, on background, off the record
Narrative style	Author-driven, reader-driven, hybrid
Type of story	By just the facts, data-driven news stories, analysis and background, deep investigations
Temporal dimension	Past, present, future
Focus	What, why, how
Purpose	Inform, persuade, entertain, explain, unknown, other
Way of communicating	Visualize, narrate, socialize, humanize, personalize, utilize
Data	
Number of data sources	
Data sources	No data sources; government or public institutions; private corporations; other organizations such as NGOs, research institutes, universitites, etc; own sources; leaks; other
Attribution	On the record, on background, off the record
Transparency level	No mention of the source of the information and the dataset; mentions the source but not the dataset; no mention of the source, but information about the data is mentioned; mentions the source and information about the data
Data access	No, yes
Methodology	No, yes
Visualizations	
Number of visualizations	
Type	Tables and lists, static graphs, interactive graphs, maps, visual resources, animations, infographics, other
Interactivity	No, yes
Interactive functions	No interactive functions, inspect, connect, select, filter, summarize, reconfigure, narrate, play, personalize, other



https://elpais.com



https://www.derstandard.at

The variables examined in the first dimension are specified below.

Topics
Based on different categorizations proposed by **Loosen**, **Reimer**, and **Schmidt** (2017), **Tandoc** and **Oh** (2017), and **Stalph** (2018), the following list of topics was created: (1) politics; (2) society; (3) economics and business; (4) health, science, and environment; (5) education; (6) sports; (7) culture and art; and (8) entertainment.

Sources
Building on prior research (**Knight**, 2015; **Loosen**; **Reimer**; **Schmidt**, 2017), we encoded the types of sources used in the analyzed news stories. We coded the number of included personal sources, the types of sources, and the attribution.

Narrative style
It was coded as follows: (a) explanatory or author-driven style, that is, a heavy message load and reduced interactivity; (b) exploratory or reader-driven style, that is, no specific order and free interaction; and (c) hybrid, that is, a combination (**Segel**; **Heer**, 2010).

Type of story
Building on the conceptualization of **Rogers** (2014) and based on the type of data and the methods employed, we distinguished between the following types of data journalism projects: (1) by just the facts (includes public data and produces a visualization), (2) data-driven news stories (refers to issues that are in the public eye and reveals the numbers behind the news), (3) analysis and background (explains the facts behind the news by providing a thorough analysis and detailed information), and (4) deep investigations (in-depth research on large datasets).

Temporal dimension, focus, purpose, way of communicating
Both the time and focus dimensions were delimited according to criteria from **Davenport** (2014). In line with **Gray**, **Bounegru**, and **Chambers** (2012), we selected the purpose of the story, while the way in which data were communicated was determined by using the classification from **Bradshaw** (2011): visualize, narrate, socialize, humanize, personalize, and utilize.

Data
Focusing on the data dimension, we also analyzed the number, the type, and the attribution of data sources (**Knight**, 2015; **Loosen**; **Reimer**; **Schmidt**, 2017). Moreover, the level of transparency in terms of access to raw data and incorporation of methodological details was also assessed (**Hayes**; **Singer**; **Ceppos**, 2007; **Karlsson**, 2010; **Parasie**; **Darigal**, 2012; **Tandoc**; **Oh**, 2017; **Young**; **Hermida**; **Fulda**, 2018; **Stalph**, 2018; **Lowrey**; **Hou**, 2018; **Zamith**, 2019; **Zang**; **Feng**, 2019).

Visualizations
Finally, building on **Córdoba-Cabús** and **García-Borrego** (2020), we coded the type of visualization as the third dimension of Table 1: (1) tables and lists, (2) static graphs, (3) interactive graphs, (4) maps with or without interactivity, (5) visual resources, (6) animations, and (7) infographics. In so doing, we also examined the actions that the audience can conduct with the visualizations, i.e., the interactivity (**Schulmeister**, 2003; **Yi**; **Ah-Kang**; **Stasko**, 2007; **Boy**, **Detiene**; **Fekete**, 2015).

Intracoder reliability
One coder coded all the news stories. To calculate the intracoder reliability, 10% of the material was coded by the same coder again 10 weeks after the first coding. Intracoder reliability scores ranged from .93 to 1 (Holsti)/.84 to 1 (Krippendorff's alpha).

> Austria was at the bottom of the global right to information ranking in 2022, with 33 points out of 150. Spain scored 73 points

2.2. Data analysis

Data were analyzed using *SPSS* (*version 25.0*). The nonparametric Mann-Whitney *U* test for two independent samples was used to analyze quantitative variables (the number of personal and data sources and visualizations). This decision was made on the basis of the sample size and the nonsatisfaction of normality (as determined by the Kolmogorov-Smirnov test). A chi-squared test (χ^2) was applied for all other variables to identify significant differences between the Spanish newspaper *El país* and the Austrian newspaper *Der standard*. Yates' continuity correction (χ_y^2) was applied when there was only one degree of freedom. If statistical significance was detected ($p \leq 0.05$), we calculated the effect size (ES) and analyzed the standardized residuals. For both, the confidence level was set at 95%.

2.3. Interviews

We conducted focused interviews with members of the data journalism teams of *El país* and *Der standard*. This technique is considered the most suitable approach to deal with innovations in this area (**Arias-Robles**; **López-López**, 2020), and it enabled us to explore the position of each media outlet. The interviewees were located and contacted through *Twitter* and corporative email. Owing to the explorative nature of this study, we talked to only one journalist per newsroom. On the basis of relevance criteria and availability, we interviewed Daniele Grasso, Coordinator of the Data Journalism Unit at *El país*, and Michael Matzenberger, Head of Data Journalism / Interactive, at *Der standard.* The interviews, lasting between 60 and 75 minutes, were conducted online via *Google Meet* between October and December 2021. The conversations were recorded and manually transcribed. A questionnaire comprising 16 questions was used, organized into three thematic areas related to the research questions: (a) the organizational structure of data teams and the common features of day-to-day data-driven stories; (b) quality data journalism, access to information, and profitability; and (c) the current status and future developments of data-driven reporting.

3. Results

3.1. Features of data-driven stories

The results of the content analysis reveal the differences between and similarities of the data journalism practices in Spain and Austria. In total, 136 data-driven stories were published during the period examined: 99 in *El país* (72.8%) and 37 in *Der standard* (27.2%). On average, the names of two journalists were mentioned per news story (ME=2; standard deviation [SD]=1.84).

In the total sample analyzed, health, science, and environment was the dominant topic (36.8%). In this context, stories that examined data about the coronavirus pandemic were especially common (Table 2). The results highlighted the significant differences between the two newspapers analyzed (ES=0.443). While *El país* published more data-driven news stories about politics (36.4%), in *Der standard*, business (24.3%) and social issues (18.3%) were the most frequently covered topics.

Next, we analyzed which sources were incorporated in data-driven news stories. The nonparametric Mann-Whitney *U* test ($U = 2.49$, $p < 0.001$) showed that the use of personal sources was significantly higher in the Austrian newspaper ($ME_{Der\ standard} = 2$; $SD = 2.22$; $ME_{El\ país} = 0$; $SD = 2.64$). *Der standard* used academic (24.3%) and business sources (24.3%) in preference over government and institutional sources (16.2%), civil society and personal testimonies of ordinary people (both 10.8%) or cultural, artistic, and sports personalities (5.4%). *El país* also relied mostly on statements by experts (20.2%), followed by civil society (12.1%), institutions (7.1%), business (5.1%), and individual testimonies (4%). The application of the Yates correction revealed that the difference between the media outlets lies in the use of quotes from business people ($\chi_y^2[1, 136] = 8.848$; $p < 0.01$; $ES = 0.282$).

Then, we considered the narrative style of the data-driven stories. The results presented in Table 3 reveal that, while in the Spanish newspaper *El país* the hybrid style (i.e., a combination of an exploratory and explanatory approach) was the most common type, in the Austrian newspaper *Der standard* the explanatory style (i.e., guiding readers through the article instead of letting them explore) was dominant (56.8%).

Table 2. Topics of the analyzed news stories

Topic	*El país*	*Der standard*	Total
Politics	36.4%	2.7%	27.2%
Society	7.1%	18.3%	10.3%
Economy and business	8.1%	24.3%	12.5%
Health, science, and environment	38.4%	32.4%	36.8%
Education	-	-	-
Sports	8.1%	13.5%	9.6%
Culture and arts	-	-	-
Entertainment	1.0%	-	0.7%
Other	1.0%	8.1%	2.9%

Note: $N = 136$; $\chi^2[6, 136] = 26.654$; $p < 0.001$

Table 3. Narrative styles used in the data-driven news stories analyzed

Narrative style	*El país*	*Der standard*	Total
Explanatory	34.3%	56.8%	40.4%
Exploratory	2.0%	–	1.5%
Hybrid	63.6%	43.2%	58.1%

Note: $N = 136$; $\chi^2[2, 136] = 6.441$; $p < 0.05$

When looking at the types of data journalism project that were reported (Table 4), *El país* and *Der standard* both focused on publishing data-based news stories (54.5% and 51.4%, respectively). They dealt with issues that were in the public eye and revealed the figures behind the latest news. At the same time, articles with exhaustive analysis and background were frequent (38.4% and 48.6%, respectively).

Table 4. Types of data journalism projects reported in the analyzed news stories

Type of story	*El país*	*Der standard*	Total
By just the facts	6.1%	–	4.4%
Data-based news stories	54.5%	51.4%	53.7%
Analysis and background	38.4%	48.6%	41.2%
Deep investigations	1.0%	–	0.7%

Note: $N = 136$; $\chi^2[3, 136] = 5.163$; n.s.

The most common timeframe used in the analyzed news stories was the present (94.1%; Table 5). A total of 69.1% of the news stories used data from the past. Less frequent were projects that include predictive analytics or references to the future (16.2%). Given this predominance of the present, the focus was explaining what happened. To a lesser extent, these articles incorporated the underlying factors that caused the result (56.6%) and how to address the issue (7.4%). *Der standard* published significantly more stories that focused on the "why" (75.7%) compared with *El país* (49.5%; $\chi_y^2[1, 136] = 6.488$; $p < 0.01$). The main purpose of these data-driven stories was to inform (100%) and, to a lesser degree, also to explain the background (68.4%). In this regard, *El país* and *Der standard* did not differ significantly.

Table 5. Timeframe used in the analyzed news stories. Note: $N = 136$

Time	*El país*	*Der standard*	Total	χ^2 test with Yates correction
Past	65.7%	78.4%	69.1%	$\chi_y^2[1, 136] = 1.490$, n.s.
Present	93.9%	94.6%	94.1%	$\chi_y^2[1, 136] = 0.000$, n.s.
Future	18.2%	10.8%	16.2%	$\chi_y^2[1, 136] = 0.604$, n.s.

The same applied to the communication style. As shown in Table 6, both newspapers used narration (98.5%) and visualization (98.5%) very often. The Spanish newspaper published visualizations in all its publications –only two news stories were not accompanied by text. On the other hand, the Austrian outlet displayed two pieces in purely textual form. Only 11% of the projects explained how the figures affected citizens, while 5.1% turned the data into personal stories.

Table 6. Overview of different ways of communicating in the analyzed news stories. Note: $N = 136$

Way of communicating	*El país*	*Der standard*	Total	χ^2 test with Yates correction
Narrate	98.0%	100%	98.5%	$\chi_y^2[1, 136] = 0.005$, n.s.
Visualize	100%	94.6%	98.5%	$\chi_y^2[1, 136] = 2.341$, n.s.
Socialize	–	–	–	–
Humanize	4.0%	8.1%	5.1%	$\chi_y^2[1, 136] = 0.270$, n.s.
Personalize	13.1%	5.4%	11.0%	$\chi_y^2[1, 136] = 0.946$, n.s.
Utilize	–	–	–	–

In both newspapers, the projects contained data from two different sources on average ($M_{total} = 2.74$; $SD = 2.42$; $M_{El\ país} = 2.98$; $SD = 2.75$; $M_{Der\ standard} = 2.19$; $SD = 1.28$). Table 7 provides an overview of the coded data sources. The main sources of the figures were governments or official institutions (72.1%), a long way above private companies (36.8%), organizations such as NGOs, and research centers or universities (32.4%), and those compiled by the newsroom itself (8.8%). It is in this last element where a significant difference between newspapers could be identified (Table 7). In 18.9% of cases, *Der standard* included numerical information collected by the members of the team through surveys, while in *El país* this was the case in 5.1% of the news stories analyzed (ES = 0.218). The attribution to these sources was mostly direct (96% and 100%, respectively).

Table 7. Overview of data sources of the figures displayed in the analyzed news stories. Note: $N = 136$

Data source	*El país*	*Der standard*	Total	χ^2 test with Yates correction
Public	74.7%	64.9%	72.1%	$\chi_y^2[1, 136] = 0.862$, n.s.
Private	41.4%	24.3%	36.8%	$\chi_y^2[1, 136] = 2.688$, n.s.
Organizations	32.3%	32.4%	32.4%	$\chi_y^2[1, 136] = 0.000$, n.s.
Own sources	5.1%	18.9%	8.8%	$\chi_y^2[1, 136] = 4.831$, $p < 0.05$
Leaks	–	–	–	–
Other	–	2.7%	0.7%	$\chi_y^2[1, 136] = 0.264$, n.s.

In general, it was possible to identify the sources in an appropriate way –directly or indirectly– but more than half of all the news stories analyzed did not provide additional information about the dataset (61%). This prevented us from determining how the figures were structured and what their features were. Only 36.8% corresponded to the norma-

tive expectations of transparency, as they incorporated both references to the origin of the figures and methodological details about the dataset. The findings presented in Table 8 indicate that *El país* showed significantly higher levels of transparency (43.3%) than *Der standard* (18.9%).

Table 8. Transparency in the analyzed news stories

Level of transparency	*El país*	*Der standard*	Total
No, source – no, methodology	3.0%	–	2.2%
Yes, source – no, methodology	53.5%	81.1%	61.0%
Yes, source – yes, methodology	43.3%	18.9%	36.8%

Note: $N = 136$; $\chi^2[2, 136] = 8.873$; $p < 0.01$

The chi-squared contrast statistic –for the level of transparency– and the adjustment made by Yates' continuity correction –for the methodology– showed the distinct nature of *El país* with a greater commitment to transparency in its publications (Table 8). However, when it comes to providing access to raw data, both newspapers provided access in only around 20% of all analyzed news stories (*El país,* 20.2% ; *Der standard,* 27%; χ_y^2 [1, 136] = 0.387, n.s.).

Moreover, we analyzed the types of visualizations applied in the data-driven stories. *El país* (ME = 4; SD = 9.47) included more visualizations compared with *Der standard* (ME = 2; SD = 1.99; $U = 751.000$, $p < 0.05$). The findings presented in Table 9 reveal that the two newspapers also differed in the types of visualization used in the news stories analyzed. While, in *El país*, interactive graphs (50.5%) were the most common type, *Der standard* preferred static graphs (48.6%).

Table 9. Type of visualizations used in the analyzed news stories. Note: $N = 136$

Type of visualization	*El país*	*Der standard*	Total	χ^2 test with Yates correction
Tables and lists	45.5%	10.8%	36.0%	$\chi_y^2[1, 136] = 12.562$, $p < 0.001$
Static graphs	38.4%	48.6%	41.2%	$\chi_y^2[1, 136] = 0.786$, n.s.
Interactive graphs	50.5%	24.3%	43.4%	$\chi_y^2[1, 136] = 6.488$, $p < 0.01$
Maps	30.3%	18.9%	27.2%	$\chi_y^2[1, 136] = 1.235$, n.s.
Graphic resources	2.0%	–	1.5%	$\chi_y^2[1, 136] = 0.005$, n.s.
Animation	3.0%	2.7%	2.9%	$\chi_y^2[1, 136] = 0.000$, n.s.
Infographics	13.1%	27.0%	16.9%	$\chi_y^2[1, 136] = 2.778$, n.s.
Other	1.0%	–	0.7%	$\chi_y^2[1, 136] = 0.000$, n.s.

Finally, we investigated which interactive functions were used by the two newspapers. The findings presented in Table 10 show that the most common functions were to inspect (51.5%), select (26.5%), and filter (23.5%).

Table 10. Interactive functions included in the analyzed news stories. Note: $N = 136$

Interactive functions	*El país*	*Der standard*	Total	χ^2 test with Yates correction
Inspect	57.6%	35.1%	51.5%	$\chi_y^2[1, 136] = 4.569$, $p < 0.05$
Connect	2.0%	–	1.5%	$\chi_y^2[1, 136] = 0.005$, n.s.
Select	25.3%	29.7%	26.5%	$\chi_y^2[1, 136] = 0.095$, n.s.
Filter	28.3%	10.8%	23.5%	$\chi_y^2[1, 136] = 3.650$, n.s.
Summarize	12.2%	5.4%	10.4%	$\chi_y^2[1, 136] = 0.716$, n.s.
Reconfigure	1.0%	–	0.7%	$\chi_y^2[1, 136] = 0.000$, n.s.
Narrate	2.0%	8.1%	3.7%	$\chi_y^2[1, 136] = 1.362$, n.s.
Play	1.0%	2.7%	1.5%	$\chi_y^2[1, 136] = 0.000$, n.s.

Following these insights into the content of the news stories analyzed, below we present the findings from the interviews with the journalists from *El país* and *Der standard.*

3.2. Organization of the teams and creation of data-driven stories

The *El país* data team is formed of three journalists with different profiles, integrated into the "Visual Narratives" group, which is made up of around ten people, including figures such as designers and developers. This team plays a cross-functional role in the newsroom, and they usually work together in all steps. The *Der standard* data unit has a similar structure, with four members with well-defined profiles. However, only the head of the team has a permanent full-time contract, which makes it difficult to create pieces on a continuous basis. As at *El país*, it has a transversal role, but there is a clear division of tasks, especially regarding technical functions such as programming, analysis, or interaction. Although the production process of data-driven stories is not standardized, both the Spanish and Austrian newspapers include four key steps:

(1) Choice and conception of the project, which consists of focusing on the topic, searching for sources of information, and previous research.

(2) Analysis of the dataset, which is the most extensive stage because the framework is established here; this includes cleaning up the figures and asking the right questions about the data.

(3) Fitting and checking the results, which refers to seeking advice and cross-checking the analysis with experts in the field.

(4) Publication, which involves selecting the suitable tool and visualizations to communicate the results.

> The coronavirus health pandemic highlighted the relevance of data journalism, as well as the rise of stories based on figures in Spain

The team spirit at the Spanish media outlet is very similar to at the Austrian one. Nevertheless, the main difference is the allocation of tasks during the production process. While in *Der standard* each member is responsible for a part of the story and these are combined at the end, at *El país* all the decisions are jointly made. *El país* devotes at least five days to publishing a data-driven story. Daniele Grasso noted,

> "We try not to tread on the daily journalism because data analysis requires time and that does not fit in with the speed of publication of a newspaper. From the beginning, we made it clear to our bosses: if they want good work, it cannot be done in one day. That does not mean that we are not able to produce articles in a morning as we have done with Covid or election results."

The *Der standard* data team ensure that they can offer immediacy because they mostly depend on other sections. To a lesser extent –especially when they propose their own topics– they publish stories with a longer production process.

For *El país,* the main obstacles to data journalism practice are the formats in which the public institutions publish their data –hindering their reuse– and the lack of availability of certain figures. In Austria, there are even greater problems. In addition to immediacy, there are also limitations related to access to information. Marzenberger explained,

> "We definitely have a real problem in Austria. The Covid pandemic showed that authorities hide data and ignore our enquiries. Even some available and anonymized data is restrained under the cloak of privacy. On top of that we have to add endless bureaucracy to get access to public databases, which makes our work more difficult and makes it not worthwhile to spend time requesting data."

3.3. Quality data journalism, access to information, and profitability

El país maintains that quality data journalism work should be interesting, useful, and transparent. Grasso stated that

> "there are no elements that determine the quality of a project; the main objective pursued by its publication is what makes it a good piece."

However, he insisted on the importance of incorporating how data is collected and how these conclusions have been reached to provide solvency to the article. In the same vein, *Der standard* also states that a good data story must comply with journalistic principles, and its quality will be determined by the news value of the information.

Both the Spanish and the Austrian newspapers agree that data journalism is more well-recognized and developed in countries with effective access-to-information laws. Although the Spanish legislation has its shortcomings, the lack of guidelines in terms of access to information in Austria makes data journalism practices more difficult.

> "[In Spain] The transparency council should have more power. It should not be an agency that only urges on paper, but it should enforce the law. That could mean a leap in quality and to duck out once questioned by the citizen would be impossible for the institutions," Grasso said.

Michael Matzenberger explained that

> "some opposition parties and civil activists have been fighting for years to bring a transparency law on track, yet we do not know if and to what extent it will become true."

> Data journalism is more well-recognized and developed in countries with effective access-to-information laws

The *El país* journalist asserted that their readers are interested in data journalism based on reading time and the online reach of their publications. This has been demonstrated with the current subscription model of this media outlet, too. He also stated that this specialization "is profitable" because it helps to add differentiation from the competition and allows them to build different and varied stories. The same is true for *Der standard*, which claimed that data-driven stories published during the pandemic are the most visited pages in their history and that these works have demonstrated the need for reliable data to make decisions. However, the team members do not feel sufficiently valued in their own newsroom.

> "Our colleagues come to us to improve their stories, but they do not see us as indispensable. It would not really matter if our team broke away," Matzenberger stated.

3.4. Insights into the state and the future of data journalism

Daniele Grasso confirmed the exponential rise of data journalism in Spain in recent years.

> "It is a type of journalism that has proved to be very useful. So useful that people think of it as good journalism and not as something specific." (Grasso)

Despite this rise, Grasso stated that Spanish data journalism suffers from a lack of recognition for two fundamental reasons: (1) the incipient origin of this specialization in the media outlets, having emerged in 2013, and (2) the late appearance of data journalism in the principal generalist newspaper and in the national public media. In Austria, data journalism is still underrepresented. Although the *Der standard* data team is formed of only four people, Michael Matzenberger argued that they are one of the largest data teams in the country. The Austrian journalist pointed to executive and management positions and the poor recognition of the potential of data journalism as the main reasons for this situation in Austria. This results in a lack of funding and resources to address data-driven stories.

> The scant presence of deep investigations during the period under review reveals the pressure for immediacy and a lack of time to deal with other issues

In this sense, the journalists at both *El país* and *Der standard* predicted that data journalism will develop constantly. Grasso forecasted that the "data journalism" label will be diluted and treated as an intrinsic part of journalism.

> "Using graphs and maps in works based on hundreds of thousands of data will become so common that there will be no need to include the data journalism label," said Grasso.

Meanwhile, according to Matzenberger, data journalism will continue to evolve. Unfortunately, unless the situation is reversed, Austria will remain at lower levels in this aspect, not in terms of education –journalism training for data work at university or schools– but on the professional side.

4. Discussion and conclusions

The aim of this study was to explore data-driven journalism in terms of the content of news stories and journalistic practices and routines. The results of this study present strong connections with previous work, but this study allows us to delve deeper into the state of data journalism in underexplored countries such as Spain and Austria. The findings from the content analysis and interviews reflect the disparity of contexts in which data journalism is carried out, the common features, the organizational structure, the limitations, and the prospects for the specialization in these countries.

The analysis reveals a general profile of the works published in *Der standard* and *El país* and highlights the similarities of and differences between these two newspapers (RQ1). Health, science, and environment is the prevailing topic. This confirms that this topic is a determining issue corresponding to the time of publication. They coincide in giving more space to articles about the coronavirus pandemic, but differ in offering space to politics –which is more represented in *El país*– and to economy or social issues –which is more represented in *Der standard*. These differences reflect the transversality of the tool, which is adaptable to any thematic area (RQ1a).

Although adding personal sources is uncommon in day-to-day data journalism, it is more frequent at the Austrian outlet. In regard to incorporating, the vast majority opt for experts in a field to complete the information. As regards narrative style, *Der standard* chose an explanatory structure (**Loosen**; **Reimer**; **Schmidt**, 2017; **Ojo**; **Heravi**, 2018), while *El país* preferred to combine explanatory and exploratory styles with some freedom for interaction (**Córdoba-Cabús**; **García-Borrego**; **López-Martín**, 2020).

The projects are almost entirely data-driven stories whose main goal is to inform and to explain what happens. The scant presence of deep investigations during the period under review reveals the pressure for immediacy and a lack of time to deal with other issues. The most common way of communicating numerical information is to combine narrative and visualization (**Loosen**; **Reimer**; **Schmidt**, 2017; **Ojo**; **Heravi**, 2018; **Young**; **Hermida**; **Fulda**, 2018; **Córdoba-Cabús**, 2020; **Rubio-Jordán**, 2021). Data journalism articles can be displayed through a variety of forms. However, it has been demonstrated that visualization plays a key role in this discipline.

The analyzed data-driven stories incorporated two data sources on average, and they mainly relied on government sources, which is in line with previous research from other countries (**Parasie**; **Darigal**, 2012; **Knight**, 2015; **De-Maeyer** *et al.*, 2015; **Cushion**; **Lewis**; **Callaghan**, 2017; **Loosen**; **Reimer**; **Schmidt**, 2017; **Stalph**, 2018; **Young**; **Hermida**; **Fulda**, 2018; **Appelgren**, 2018; **Hyder**; **Nahid**, 2019). The problems regarding access to public information in Austria are reflected in the use of the news outlets' own sources. Unlike *El país*, *Der standard* relied more on its own surveys and experiments to produce its own data for its news stories. This strategy allowed it to overcome the opacity of the Austrian administration.

Although transparency is considered to be a fundamental element in data journalism stories, this characteristic is scarce in day-to-day articles according to previous studies (**Lowrey**; **Hou**, 2018; **Young**; **Hermida**; **Fulda**, 2018; **Stalph**, 2018; **Zhang**; **Feng**, 2019; **Zamith**, 2019; **Córdoba-Cabús**; **García-Borrego**, 2021). While the vast majority of articles cite sources directly, linking to datasets and adding a methodological section remains rare. **Zamith** (2019, p. 15) stated that

> "journalists may not prioritize leading readers to source material when attempting to meet deadlines during their day-to-day work or actively resist ceding too much autonomy through over-sharing."

> Data journalism articles can be displayed through a variety of forms. However, it has been demonstrated that visualization plays a key role in this discipline

In this sense, *El país* differs from *Der standard*. The Spanish outlet meets the idealized form of transparency on more occasions (RQ1b). This may be due to the lack of resources, time, and knowledge to provide such openness (**Córdoba-Cabús**; **García-Borrego**, 2021) or, as mentioned by **Porlezza** and **Splendore** (2019, p. 1246), "because of the fear of free-riding by the competitors," which could involve copying or replicating their work.

> The projects increasingly employ perceived interactivity, which gave the audience a sense of control, while the author actually retained full control

The strategy in terms of visualizations and interactivity also differs (RQ1c). While *El país* added interactive graphs, *Der standard* opted for static graphs. Despite these divergences, the findings reinforce the idea of reduced interactivity (**Loosen**; **Reimer**; **Schmidt**, 2017; **Young**; **Hermida**; **Fulda**, 2018; **Stalph**, 2018; **Ojo**; **Heravi**, 2018; **Appelgren**, 2018; **Córdoba-Cabús**; **García-Borrego**, 2020). This decrease in interactivity is closely aligned with the addition of simpler and more basic interactive features in visualizations, such as exploration. As pointed out by **Appelgren** (2018),

> "The illusion of interactivity thus replaces real interactivity."

The projects increasingly employ perceived interactivity, which gave the audience a sense of control, while the author actually retained full control.

The interviews confirmed that the context strongly conditions how data journalism is carried out (**Appelgren**; **Nygren**, 2014; **Appelgren**; **Salaverría**, 2018; **Sandoval-Martín**; **La-Rosa**, 2018; **Porlezza**; **Splendore**, 2019). On the one hand, the incipient status of data journalism in Austria is mainly due to the lack of guidelines in terms of access to information (RQ2). Moreover, there is an absence of awareness of their value, which implies a lack of recognition even within their own newspaper. As **Tong** (2022, p. 2795) stated,

> "Data journalists' expertise, together with the increasing internal recognition and their compelling data stories, helps them establish cultural authority."

On the other hand, at *El país*, the working schedule is respected and data journalism is used as a feature to differentiate it from the competition. The articles produced during the analyzed period reveal the difference in this regard (99 news stories published in *El país* and 37 in *Der standard*). The situation of Austrian data journalism is more similar to that in Arab countries (**Lewis**; **Nashmi**, 2019; **Mutsvairo**; **Bebawi**; **Borges-Rey**, 2020) than to that in European countries. Austrian data journalists face all the obstacles mentioned in the scientific literature, both internal –within the newsrooms– and external –related to the context (**Appelgren**; **Lindén**; **Van-Dalen**, 2019; **Kashyap**; **Bhaskaran**; **Mishra**, 2020).

The data teams –based on the number of workers– and the production process of a data journalism story are similar in *El país* and *Der standard*. Although the inverted pyramid proposed by **Bradshaw** (2011) is called into question (**Albert-Trinidad**, 2020), these findings show that data-driven articles are produced according to these phases. The differences lie in the organizational structure of the data teams and in the contractual relationships with each company. Data journalists in *El país* work together in all steps and have permanent full-time contracts. However, a division of tasks prevails in *Der standard*, and the members of the data journalism team are hired part-time.

The perceptions about the future of data journalism are in line with **Stalph** and **Borges-Rey** (2018) and **Ëkström** and **Westlund** (2019): Data will become an inherent part of journalism, and it will be essential to all journalistic publications (RQ3). Journalists at both *El país* and *Der standard* believe that data journalism will grow and evolve. Nevertheless, obstacles in Austria are expected to continue to limit the practice of data journalism.

This study is not free of limitations. First, our sample for the content analysis was limited. We were only able to analyze one newspaper per country. Hence, future studies should extend such analysis to other newspapers that practice data journalism in different countries. In addition, it would be interesting to analyze changes over time and to deepen our understanding of the impact of data journalism on society.

Notwithstanding these limitations, this study was able to deliver important initial insights into day-to-day data journalism content and work routines in Spain and Austria. The practice of data journalism in Austria and Spain mainly differs because of the context in which it is carried out. This research confirms the relevance that this specialization is currently acquiring in countries such as Spain and demonstrates the challenges that journalists face in countries without access to information and transparency laws to build good stories with data.

> Journalists face challenges to build good stories with data in countries without access to information and transparency laws

5. References

Albarran, Alan (2009). *The handbook of Spanish language media*. New York: Routledge. ISBN: 978 0 415991018

Albert-Trinidad, Ramaris (2020). "Data journalism in the Spanish Caribbean digital media". *Anàlisi*, num. extra, pp. 47-66. *https://doi.org/10.5565/rev.analisi.3236*

Anderson, Christopher-Williams (2015). "Between the unique and the pattern: Historical tensions in our understanding of quantitative journalism". *Digital journalism*, v. 3, n. 3, pp. 349-363.
https://doi.org/10.1080/21670811.2014.976407

Appelgren, Ester (2018). "Openness and transparency in the data journalism practice". *Nordicom information*, v. 40, n. 1, pp. 76-77.
https://www.nordicom.gu.se/sites/default/files/kapitel-pdf/appelgren.pdf

Appelgren, Ester; **Lindén, Carl-Gustav** (2020). "Data journalism as a service: Digital native data journalism expertise and product development". *Media and communication*, v. 8, n. 2, pp. 62-72.
https://doi.org/10.17645/mac.v8i2.2757

Appelgren, Ester; **Lindén, Carl-Gustav**; **Van-Dalen, Arjen** (2019). "Data journalism research: Studying a maturing field across journalistic cultures, media markets and political environments". *Digital journalism*, v. 7, n. 9, pp. 1191-1199.
https://doi.org/10.1080/21670811.2019.1685899

Appelgren, Ester; **Nygren, Gunnar** (2014). "Data journalism in Sweden: Introducing new methods and genres of journalism into 'old' organizations". *Digital journalism*, v. 2, n. 3, pp. 394-405.
https://doi.org/10.1080/21670811.2014.884344

Appelgren, Ester; **Salaverría, Ramón** (2018). "The promise of the transparency culture: A comparative study of access to public data in Spanish and Swedish newsrooms". *Journalism practice*, v. 12, n. 8, pp. 986-996.
https://doi.org/10.1080/17512786.2018.1511823

Arias-Robles, Félix; **Carvajal, Miguel** (2022). "Data journalists. Analysis of the sociodemographic profile of the professional in Spain and Latin America". *Estudios sobre el mensaje periodístico*, v. 28, n. 1, pp. 621-637.
https://doi.org/10.5209/esmp.76395

Arias-Robles, Félix; **López-López, Pilar-José** (2020). "Driving the closest information. local data journalism in the UK". *Journalism practice*, v. 15, n. 5, pp. 638-650.
https://doi.org/10.1080/17512786.2020.1749109

Arrese, Ángel (2022). "In the beginning were the data: Economic journalism as/and data journalism". *Journalism studies*, v. 23, n. 4, pp. 487-505.
https://doi.org/10.1080/1461670X.2022.2032803

Ausserhofer, Julian; **Gutounig, Robert**; **Oppermann, Michael**; **Matiasek, Sarah**; **Goldgruber, Eva** (2017). "The datafication of data journalism scholarship: Focal points, methods, and research propositions for the investigation of data-intensive newswork". *Journalism*, v. 21, n. 7, pp. 950-973.
https://doi.org/10.1177/1464884917700667

Borges-Rey, Eddy (2016). "Unravelling data journalism: A study of data journalism practice in British newsrooms". *Journalism practice*, v. 10, n. 7, pp. 833-843.
https://doi.org/10.1080/17512786.2016.1159921

Boy, Jeremy; **Détienne, Françoise**; **Fekete, Jean-Daniel** (2015). "Storytelling in information visualizations: Does it engage users to explore data?". In: *Proceedings of the 33rd annual ACM conference on human factors in computing systems*.
https://doi.org/10.1145/2702123.2702452

Bradshaw, Paul (2011). *6 ways of communicating data journalism*.
https://bit.ly/3dbcjDl

Camaj, Lindita; **Martin, Jason**; **Lanosga, Gerry** (2022). "The impact of public transparency infrastructure on data journalism: A comparative analysis between information-rich and information-poor Countries". *Digital journalism*, online first.
https://doi.org/10.1080/21670811.2022.2077786

Chaparro-Domínguez, María-Ángeles; **Díaz-Campo, Jesús** (2021). "Data journalism and ethics: Best practices in the winning projects (DJA, OJA and Sigma Awards)". *Journalism practice*, online first.
https://doi.org/10.1080/17512786.2021.1981773

Cheruiyot, David; **Baack, Stefan**; **Ferrer-Conill, Raúl** (2019). "Data journalism beyond legacy media: The case of African and European civic technology organizations". *Digital journalism*, v. 7, n. 9, pp. 1215-1229.
https://doi.org/10.1080/21670811.2019.1591166

Coddington, Mark (2015). "Clarifying journalism's quantitative turn. A typology for evaluating data journalism, computational journalism and computer-assisted reporting". *Digital journalism*, v. 3, n. 3, pp. 331-348.
https://doi.org/10.1080/21670811.2014.976400

Córdoba-Cabús, Alba (2018). "Análisis del periodismo de datos en la campaña electoral del 20D a través de las ediciones digitales de diarios generalistas". *Estudios sobre el mensaje periodístico*, v. 24, n. 1, pp. 137-154.
https://doi.org/10.5209/esmp.59942

Córdoba-Cabús, Alba (2020). "Estándares de calidad en el periodismo de datos: fuentes, narrativas y visualizaciones en los Data Journalism Awards 2019". *Profesional de la información*, v. 29, n. 3, e290328.
https://doi.org/10.3145/epi.2020.may.28

Córdoba-Cabús, Alba; **García-Borrego, Manuel** (2020). "Visualisations as a critical information source for data journalism. Analysis of the typology, interactivity, and functions in the 2019 Data Journalism Awards". *Doxa comunicación*, n. 31, pp. 87-105.
https://doi.org/10.31921/doxacom.n31a4

Córdoba-Cabús, Alba; **García-Borrego, Manuel** (2021). "Evolución de la transparencia en el periodismo de datos de referencia. Estudio de las historias publicadas entre 2018 y 2019". *Icono14*, v. 19, n. 2, pp. 364-387.
https://doi.org/10.7195/ri14.v19i2.1648

Córdoba-Cabús, Alba; **García-Borrego, Manuel**; **López-Martín, Álvaro** (2020). "El periodismo de datos durante la crisis sanitaria del Covid-19 en la prensa española". *RISTI*, E35, pp. 325-337.
http://www.risti.xyz/issues/ristie35.pdf

Cortés-del-Alamo, Helena; **Luengo, María**; **Elías, Carlos** (2018). "Periodismo de datos y transparencia al margen de los grandes medios: Un estudio comparativo de Civio y ProPublica". *Icono14*, v. 16, n. 2, pp. 66-87.
https://doi.org/10.7195/ri14.v16i2.1177

Costa-Sánchez, Carmen; **López-García, Xosé** (2020). "Comunicación y crisis del coronavirus en España. Primeras lecciones". *Profesional de la información*, v. 29, n. 3, e290304.
https://doi.org/10.3145/epi.2020.may.04

Cushion, Stephen; **Lewis, Justin**; **Callaghan, Robert** (2017). "Data journalism, impartiality and statistical claims". *Journalism practice*, v. 11, n. 10, pp. 1198-1215.
https://doi.org/10.1080/17512786.2016.1256789

Davenport, Thomas (2014). *10 kinds of stories to tell with data*.
https://hbr.org/2014/05/10-kinds-of-stories-to-tell-with-data

De-Lima-Santos, Mathias-Felipe; **Schapals, Aljosha-Karim**; **Bruns, Axel** (2020). "Out-of-the-box versus in-house tools: how are they affecting data journalism in Australia?". *Media international Australia*, v. 181, n. 1, pp. 152-166.
https://doi.org/10.1177/1329878X20961569

De-Lima-Santos, Mathias-Felipe; **Mesquita, Lucia** (2021). "The strategic value of data journalism". In: Salaverría, Ramón; De-Lima-Santos, Mathias-Felipe (eds.). *Journalism, data and technology in Latin America.* London: Palgrave Macmillan, pp. 97-136. ISBN: 978 3 030 65860 1

De-Maeyer, Juliette; **Libert, Manon**; **Domingo, David**; **Heinderyckx, François**; **Le-Cam, Florence** (2015). "Waiting for data journalism: A qualitative assessment of the anecdotal take-up of data journalism in French-speaking Belgium". *Digital journalism*, v. 3, n. 3, pp. 432-446.
https://doi.org/10.1080/21670811.2014.976415

Ëkström, Mats; **Westlund, Oscar** (2019). "Epistemology and journalism". *Journalism studies*.
https://doi.org/10.1093/acrefore/9780190228613.013.806

Figl, Bettina (2017). *Bigger is not always better: what we can learn about data journalism from small newsrooms*.
https://reutersinstitute.politics.ox.ac.uk/our-research/bigger-not-always-better-what-we-can-learn-about-data-journalism-small-newsrooms

Fink, Katherine; **Anderson, Christopher Williams** (2015). "Data journalism in the United States: Beyond the usual suspects". *Journalism studies*, v. 16, n. 4, pp. 467-481.
https://doi.org/10.1080/1461670X.2014.939852

Gray, Jonathan; **Bounegru, Liliana**; **Chambers, Lucy** (2012). *The data journalism handbook*. Sebastopol: European Journalism Centre; O'Reilly Media. ISBN: 978 1 449 33006 4

Grünangerl, Manuela; **Trappel, Josef**; **Tomaz, Tales** (2021). "Austria: Confirmed democratic performance while slowly digitalizing". In: Trappel, Josef; Tomaz, Tales (eds.). *The media for democracy monitor 2021: How leading news media survive digital transformation.* Gothenburg: University of Gothenburg, pp. 95-152.
https://doi.org/10.48335/9789188855404-3

Hallin, Daniel; **Mancini, Paolo** (2004). *Comparing media systems: Three models of media and politics*. New York: Cambridge University Press. ISBN: 0 521 83535 6

Hayes, Arthur; **Singer, Jane**; **Ceppos, Jerry** (2007). "Shifting roles, enduring values: the credible journalist in a digital age". *Journal of mass media ethics*, v. 22, n. 4, pp. 79-262.
https://doi.org/10.1080/08900520701583545

Heravi, Bahareh (2018). "3WS of data journalism education". *Journalism practice*, v. 13, n. 3, pp. 349-366.
https://doi.org/10.1080/17512786.2018.1463167

Heravi, Bahareh; **Lorenz, Mirko** (2020). "Data journalism practices globally: Skills, education, opportunities, and values". *Journalism and media*, v. 1, n. 1, pp. 26-40.
https://doi.org/10.3390/journalmedia1010003

Holovaty, Adrian (2006). "A fundamental way newspaper sites need to change". *Holovaty.com*, 6th September.
http://www.holovaty.com/writing/fundamental-change

Howard, Alexander-Benjamin (2014). *The art and science of data-driven journalism*. Tow Center for Digital Journalism, Columbia Journalism School.
https://doi.org/10.7916/D8Q531V1

Hyder, Mohammed; **Nahid, Ibrahim** (2019). "Press data and information graphics, prospects and challenges of development in the Arab World". *American journal of computer science and technology*, v. 2, n. 4, pp. 48-51.
https://doi.org/10.11648/j.ajcst.20190204.11

Igartua, Juan-José (2006). *Métodos cuantitativos de investigación en comunicación*. Barcelona: Bosh. ISBN: 978 84 97902717

Karlsson, Michael (2010). "Rituals of transparency: Evaluating online news outlets' uses of transparency rituals in the United States, United Kingdom and Sweden". *Journalism studies*, v. 11, n. 4, pp. 535-545.
https://doi.org/10.1080/14616701003638400

Kashyap, Geeta; **Bhaskaran, Harikrishnan**; **Mishra, Harsh** (2020). "We need to find a revenue model: Data journalists' perceptions on the challenges of practicing data journalism in India". *Observatorio*, v. 14, n. 2, pp. 121-136.
https://doi.org/10.15847/obsOBS14220201552

Kashyap, Geeta; **Mishra, Harsh**; **Bhaskaran, Harikrishnan** (2022). "Data journalists' perception and practice of transparency and interactivity in Indian newsrooms". *Media Asia*.
https://doi.org/10.1080/01296612.2022.2076324

Knight, Megan (2015). "Data journalism in the UK: a preliminary analysis of form and content". *Journal of media practice*, v. 16, n. 1, pp. 55-72.
https://doi.org/10.1080/14682753.2015.1015801

Lewis, Norman; **Nashmi, Eisa** (2019). "Data journalism in the Arab Region: Role conflict exposed". *Digital journalism*, v. 7, n. 9, pp. 1200-1214.
https://doi.org/10.1080/21670811.2019.1617041

Loosen, Wiebke; **Reimer, Julius**; **Schmidt, Fenja** (2017). "Data-driven reporting: An on-going (r)evolution? An analysis of projects nominated for the Data Journalism Awards 2013-2016". *Journalism: theory, practice & criticism*, v. 21, n. 9, pp. 1246-1263.
https://doi.org/10.1177/1464884917735691

Lowrey, Wilson; **Hou, Jue** (2018). "All forest, no trees? Data journalism and the construction of abstract categories". *Journalism*, v. 22, n. 1, pp. 35-51.
https://doi.org/10.1177/1464884918767577

Mayer-Schönberger, Viktor; **Cukier, Kenneth** (2013). "Big data: A revolution that will transform how we live, work, and think". *American journal of epidemiology*, v. 179, n. 9, pp. 1143-1144. ISBN: 978 0 544227750
https://doi.org/10.1093/aje/kwu085

Mutsvairo, Bruce; **Bebawi, Saba**; **Borges-Rey, Eddy** (2020). *Data journalism in the Global South*. Switzerland: Palgrave Macmillan. ISBN: 978 3 030251772

Ohlsson, Jonas (2015). "The Nordic media market 2015". *Nordic research on media and communication*, v. 36, n. 1, pp. 157-158. ISBN: 978 91 87957 05 5

Ojo, Adegboyega; **Heravi, Bahareh** (2018). "Patterns in award winning data storytelling: Story types, enabling tools and competences". *Digital journalism*, v. 6, n. 6, pp. 693-718.
https://doi.org/10.1080/21670811.2017.1403291

Parasie, Sylvain; **Dagiral, Eric** (2012). "Data-driven journalism and public good: 'Computer-assisted-reporters' and 'programmer-journalists' in Chicago". *New media & society*, v. 15, n. 6, pp. 853-871.
https://doi.org/10.1177/1461444812463345

Pereira, Fabio-Henrique; **Mastrella, Bruna** (2022). "Practice of data-driven journalism in Brazil: Between disruption and the reinstatement of professional values". *Digital journalism*, v. 10, n. 4, pp. 599-625.
https://doi.org/10.1080/21670811.2022.2061550

Porlezza, Colin (2019). "Data journalism and the ethics of open source: Transparency and participation as a prerequisite for serving the public good". In: Daly, Angela; Devitt, Kate; Mann, Monique (eds.). *Good data. Theory on demand*. Amsterdam: Institute of Network Cultures, pp. 189-201. ISBN: 978 94 92302 27 4

Porlezza, Colin; **Splendore, Sergio** (2019). "From open journalism to closed data: Data journalism in Italy". *Digital journalism*, v. 7, n. 9, pp. 1230-1252.
https://doi.org/10.1080/21670811.2019.1657778

Reuters Institute (2021). *Digital News Report*.
https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2021-06/Digital_News_Report_2021_FINAL.pdf

Rogers, Simon (2014). "Data journalism is the new punk". *British journalism review*, v. 25, n. 2, pp. 31-34.
https://doi.org/10.1177/0956474814538181

Rojas-Torrijos, José-Luis; **Rivera-Hernández, Agustín** (2016). "El español y El confidencial, exponentes del periodismo deportivo en los medios nativos digitales españoles". *Doxa comunicación*, n. 23, pp. 171-193.
https://revistascientificas.uspceu.com/doxacomunicacion/article/view/554

RTI (2022). *Global right to information rating map*.
https://www.rti-rating.org

Rubio-Jordán, Ana-Virginia (2021). "Periodismo de datos y prensa nativa digital: cobertura de la campaña electoral del 10N a través de El confidencial, Eldiario.es y El español". *Doxa comunicación*, n. 32, pp. 289-303.
https://doi.org/10.31921/doxacom.n32a14

Salaverría, Ramón; **Gómez-Baceiredo, Beatriz** (2017). *Spain. Media landscapes*.
https://medialandscapes.org/country/spain

Sanahuja, Rosana; **López, Pablo** (2021). "Journalistic applications of artificial intelligence. Concept map, professional roles and developing trends in the context of the global Covid-19 pandemic". *Razón y palabra*, v. 25, n. 112, pp. 432-449.
https://doi.org/10.26807/rp.v25i112.1827

Sánchez-Calero, María-Luisa; **Mancinas-Chávez, Rosalba** (2017). "Periodismo con datos no es igual que periodismo de datos". *TecCom studies*, n. 8, pp. 53-69.
http://www.teccomstudies.com/index.php?journal=teccomstudies&page=article&op=view&path%5B%5D=49

Sandoval-Martín, Teresa; **La-Rosa, Leonardo** (2018). "Big data as a differentiating sociocultural element of data journalism: the perception of data journalists and experts". *Communication and society*, v. 31, n. 4, pp. 193-209.
https://doi.org/10.15581/003.31.4.193-209

Schumelmeister, Rolf (2003). "Taxonomy of multimedia component interactivity: a contribution to the current metadata debate". *SComS: New media in education*, pp. 61-80.
https://bit.ly/2VBfGyc

Segel, Edward; **Heer, Jeffrey** (2010). "Narrative visualization: Telling stories with data". *IEEE Transactions on visualization and computer graphics*, v. 16, n. 6, pp. 1139-1148.
https://doi.org/10.1109/TVCG.2010.179

Splendore, Sergio; **Di-Salvo, Philip**; **Eberwein, Tobias**; **Groenhart, Harmen**; **Kus, Michal**; **Porlezza, Colin** (2016). "Educational strategies in data journalism: A comparative study of six European countries". *Journalism*, v. 17, n. 1, pp. 138-152.
https://doi.org/10.1177/1464884915612683

Stalph, Florian (2018). "Classifying data journalism: A content analysis of daily data-driven stories". *Journalism practice*, v. 12, n. 10, pp. 1332-1350.
https://doi.org/10.1080/17512786.2017.1386583

Stalph, Florian; **Borges-Rey, Eddy** (2018). "Data journalism sustainability: An outlook on the future of data-driven reporting". *Digital journalism*, v. 6, n. 8, pp. 1078-1089.
https://doi.org/10.1080/21670811.2018.1503060

Tabary, Constance; **Provost, Anne-Marie**; **Trottier, Alexandre** (2016). "Data journalism's actors, practices and skills: A case study from Quebec". *Journalism*, v. 17, n. 1, pp. 66-84.
https://doi.org/10.1177/1464884915593245

Tandoc, Edson; **Oh, Soo-kwang** (2017). "Small departures, big continuities?: Norms, values, and routines in The Guardian's big data journalism". *Journalism studies*, v. 18, n. 8, pp. 997-1015.
https://doi.org/10.1080/1461670X.2015.1104260

Teruel-Rodríguez, Laura; **Blanco-Castilla, Elena** (2016). "El periodismo de datos en los medios de referencia españoles". *Trípodos*, n. 38, pp. 13-30.
http://www.tripodos.com/index.php/Facultat_Comunicacio_Blanquerna/article/view/322

Tong, Jingrong (2021). "Paradigm reinforcing: The assimilation of data journalism in the United Kingdom". *Journal of applied journalism & media studies*.
https://doi.org/10.1386/ajms_00043_1

Tong, Jingrong (2022). "Rethinking the expertise of data journalists: A case study". *International journal of communication*, n. 16, pp. 2781-2800.
https://ijoc.org/index.php/ijoc/article/view/18008

Tong, Jingrong; **Zuo, Landong** (2021). "The inapplicability of objectivity: Understanding the work of data journalism". *Journalism practice*, v. 15, n. 2, pp. 153-169.
https://doi.org/10.1080/17512786.2019.1698974

Toonders, Joris (2014). "Data is the new oil of the digital economy". *Wired*, 7th July.
https://cutt.ly/oX8nEjQ

Uskali, Turo; **Kuutti, Heikki** (2015). "Models and streams of data journalism". *The journal of media innovations*, v. 2, n. 1, pp. 77-88.
https://doi.org/10.5617/jmi.v2i1.882

Van-Dijck, Jose (2014). "Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology". *Surveillance & society*, v. 12, n. 2, pp. 197-208.
https://doi.org/10.24908/ss.v12i2.4776

Veglis, Andreas; **Bratsas, Charalampos** (2017). "Reporters in the age of data journalism". *Journalism & media studies*, v. 6, n. 2, pp. 225-244.
https://doi.org/10.1386/ajms.6.2.225_1

Wright, Kate; **Zamith, Rodrigo**; **Bebawi, Saba** (2019). Data journalism beyond majority world countries: Challenges and opportunities. *Digital journalism*, v. 7, n. 9, pp. 1295-1302.
https://doi.org/10.1080/21670811.2019.1702472

Wu, Shangyuan (2022). "Evaluating 'exemplary data journalism' from Asia: An exploration into South China Morning Post's data stories on China and the world". *Journalism*.
https://doi.org/10.1177/14648849221093509

Yi, Ji-Soo; **Ah-Kang, Youn**; **Stasko, John** (2007). "Toward a deeper understanding of the role of interaction visualization". *IEEE Transactions on visualization and computer graphics*, v. 13, n. 6, pp. 1224-1231.

Young, Mary-Lynn; **Hermida, Alfred**; **Fulda, Johanna** (2018). "What makes for great data journalism?: A content analysis of data journalism awards finalists 2012-2015". *Journalism practice*, v. 12, n. 1, pp. 115-135.
https://doi.org/10.1080/17512786.2016.1270171

Zamith, Rodrigo (2019). "Transparency, interactivity, diversity, and information provenance in everyday data journalism". *Digital journalism*, v. 7, n. 4, pp. 470-489.
https://doi.org/10.1080/21670811.2018.1554409

Zhang, Shuling; **Feng, Jieyun** (2019). "A step forward? Exploring the diffusion of data journalism as journalistic innovations in China". *Journalism studies*, v. 20, n. 9, pp. 1281-1300.
https://doi.org/10.1080/1461670X.2018.1513814