

Research Article

Data Journalism as an innovation in social communication: The case in sports industry

El periodismo de datos como innovación en la comunicación social: El caso de la industria del deporte

Zeliha Işıl Vural^{1*} and Pere Masip¹

Ramon Llull University, Blanquerna, Passeig de Sant Gervasi, 47, 08022 Barcelona.

*Correspondence: zelihaisily@blanquerna.url.edu

Abstract: Data analysis has always been an integral part of journalism but combining it with technology was a novelty for newspapers. Journalism's combination with technology was an innovation because of processing, interpretation, and visualization of large datasets in a journalistic content. In recent years, newspapers have started to adapt data journalism and integrated it to sports for better storytelling and making sports more understandable for readers. This research aims to analyse sports data journalism practices in Spain with a quantitative approach with content analysis of 1068 data journalism articles published by 6 newspapers (Marca, Mundo Deportivo, AS, El Mundo, El Periódico, El País) between 2017-2019. Quantitative analysis focuses on how sports data journalism is being adapted in Spain, technical features of articles, and the similarities and differences between sports and national newspapers to identify integration of sports data journalism.

Keywords: data journalism; digital journalism; sports journalism; data visualisation; content analysis; innovation; Spain.

Resumen: El análisis de datos siempre ha sido una parte integral del periodismo, pero combinarlo con la tecnología ha sido una novedad para la prensa escrita. La combinación del periodismo con la tecnología se ha presentado como una innovación reciente debido al procesamiento, la interpretación y la visualización de grandes conjuntos de datos utilizados en el contenido periodístico. En los últimos años, los periódicos han adaptado el periodismo de datos integrándolo, por ejemplo, en industrias como el deporte, mejorando la narración y haciendo que los deportes sean más comprensibles para los lectores. En este sentido, esta investigación tiene como objetivo analizar las prácticas de periodismo de datos deportivo en España con un enfoque cuantitativo a través del análisis de contenido de 1.068 artículos de periodismo de datos publicados por 6 periódicos (Marca, Mundo Deportivo, AS, El Mundo, El Periódico, El País) entre 2017-2019. El análisis se centra en cómo se está adaptando el periodismo de datos deportivo en España, las características técnicas de los artículos y las similitudes y diferencias entre los periódicos deportivos nacionales en relación con la integración del periodismo de datos.

Palabras clave: periodismo de datos; periodismo digital; periodismo deportivo; visualización de datos; análisis de contenidos; innovación; España.

1. Introduction

The data-driven transformation of journalism has gained momentum in recent years, and it has started to displace conventional journalism in favour of digital media. While data is used in conventional journalism, data journalism has emerged because of the digital transformation of data and its dissemination to the reader. Journalism has been reshaped several times across history. In time, journalistic methods engaged with new fields such as precision journalism and computer-assisted reporting, and the process is still ongoing. Progress in data-driven journalism began with Philip Meyer's Precision Journalism in the 1970s, and major advancements occurred after the 1980s with the emergence of computer-assisted reporting (Meyer, 1991). Many application methods for journalistic practices in order to collect, analyse, and combine data by computer processing are used in precision journalism. Meyer (1991) also refers to this journalistic activity as 'database journalism'. Based on Meyer (1991), a journalist should know "how to find information", "how to evaluate and analyse it" and "how to communicate it in a way that will pierce the babble of information overload and reach the people who need and want it". Precision journalism took a scientific approach to the techniques used and the common goals. It included both a science and a market component. In the light of these, Philip Meyer also recognized as the pioneer of Computer-Assisted Reporting (CAR). In the forthcoming years, Bruce Garrison (1998) explained the connection between precision journalism and CAR as: "Precision journalism, the code word for a generation of computer assisted reporting, but the two terms essentially describe the same quantitative approaches to news reporting". CAR changed the way journalists work and thought about news, and that transition is not solely motivated by modern technology but must be considered in a wider context (Hammond, 2015). Newsrooms became more familiar with emerging technology, analytics, visualization, and investigative journalism as a result of the use of CAR techniques in stories, and CAR developed into data-driven practices, which would be dubbed data journalism in the coming years.

2. Emergence of data journalism

Working with data has always been a crucial element of journalism but merging it with technology was an innovation. Journalists had traditional skills, but the innovation was making journalists both graphic designers and coders. Data searching, processing, interpretation, and visualisation were the new skills for journalists and data journalism was transforming the ideal journalism term to a new and innovative term. The precursors of data journalism consistently pointed to the journalists as still journalists but with data analysis and data visualisation talents. Journalists and academics made various descriptions of data journalism, but the common point of all descriptions was the using data to tell a story in a better way. However, the main description of data journalism has been made by Howard (2014) as "gathering, cleaning, organizing, analysing, visualising, and publishing data to support the creation of acts of journalism". Later, academics described data journalism based on Howard's words. Coddington (2015) pointed to an overlap of journalistic practices and highlighted the specific features of CAR, data journalism, and computational journalism. According to Coddington (2015) CAR was focusing on investigative reporting, while data journalism was mainly based on open data, and computational journalism was aiming to combine programming and algorithms. Veglis and Bratsas (2017) explained data journalism in a similar way and described it as finding data and extracting the necessary information from that data to write an article with visualisations. It was a way to enrich the article and highlight the significance of the story for readers.

When Wikileaks released the Afghan War Logs in 2010, it sparked global attention, and it has since come to be regarded as the starting point of data-driven journalism. Wikileaks, on the other hand, should not be regarded as the very first example of data journalism in terms of the stages of a data-driven report, but rather as an example of data journalism's roots as Baack (2011) claims. After Wikileaks, data journalism practices started to be seen in the US and UK. Simon

Rogers and Steve Doig integrated data journalism to newsrooms and played a crucial role in spreading data journalism (Grassler, 2018). Through the Guardian and the New York Times, data journalism practices were accepted as a profession in the newsrooms and influenced other newspapers to create data departments and publish data journalism articles (Olivo, 2015; Howard, 2014; Tandoc Jr. & Oh, 2017). Following years witnessed to the Upshot (Olivo, 2015; Howard, 2014) which is a data journalism blog by the New York Times, and the data journalism projects of the Washington Post (Craig & Zion, 2014; Schulze, 2015; Michalczki, 2016) the Financial Times, and BBC (Borges-Rey, 2016). Data journalism practices in the US and UK were followed by European countries such as Sweden (Appelgren & Nygren, 2014; Appelgren & Salaverria, 2018), Belgium (De Maeyer et al., 2015; Grassler, 2018), Greece (Charbonneaux & Gkouskou-Giannakou, 2015), Italy (Porlezza & Splendore, 2019), and Spain (Ferrerias Rodríguez, 2016; Blanco Castilla & Quesada, 2016; Rivera Hernandez & Rojas Torrijos, 2016; La-Rosa & Sandoval-Martin, 2016; Appelgren & Salaverria, 2018; Grassler, 2018; Arias et al., 2018) and data journalism became a profession in the whole world.

3. Data journalism meets sports industry

Sports and data journalism are a great match because sports have an abundance of data in every branch to help audiences, media, and industry professionals understand and interpret the game. However, there is still a complicated relationship between sports and data in terms of collecting data and developing the skills necessary to properly analyse and visualize it. Data journalist John Burn-Murdoch explained the importance of sports data journalism as: "During an event like the Olympics, this need becomes more acute because many people follow them through the internet to keep track of the dozens of events held there each day. A clear graph is a great opportunity to make the information stand out within such a large amount of data, and it can be fully understood in the same moment" (Rivera Hernandez & Rojas Torrijos, 2016). Because of the growing success of sports organizations on both a national and international level, the first sports data journalism practices began in English-speaking media with baseball and then expanded to basketball, football, and more recently other branches (Rojas Torrijos, 2016). During the Winter Olympics and World Cup in 2014, the New York Times, the Guardian, the Economics, and the Washington Post published the first examples of sports data journalism. These events provided an opportunity for newsrooms to launch new innovations and data teams during the same year (Rivera Hernandez & Rojas Torrijos, 2016). The following years saw the rise of FiveThirtyEight and Grantland, both of which play important roles in sports data journalism. Nate Silver created FiveThirtyEight as a personal project when the website was owned by the New York Times, then ESPN, and then ABC News, and won numerous awards (Rojas Torrijos, 2015). Through the projects of American and British media European countries were inspired by successful examples and began to pay more attention to sports data journalism.

4. Spanish data journalism and its integration with sports

In 2011, data journalism and various forms of data-driven projects grew in popularity in Spanish media. RTVE opened a data section under the Audiovisual Innovation Laboratory, and they released their first project, which is about the campaign pledges of Spanish political parties (Ferrerias Rodríguez, 2016). After 2011, MediaLab Prado began hosting data journalism conferences and seminars to promote data journalism. With the assistance of The International Consortium of Investigative Journalists (ICIJ), Open Knowledge Foundation (OKFN), and European Journalism Center (EJC), MediaLab Prado and CCCBLab have begun to provide some grants for data journalism and later Medialab Prado organized 'Data Journalism Seminar' in 2011 before the transparency law. This seminar is accepted as the first data journalism event in Spain (Blanco Castilla & Quesada, 2016).

Data-driven sports journalism began in 2011 with 'munideporte.com' an online sports publication and continued with 'Quality Sport' a tablet publication focused primarily on women's football in 2012 (Labarga-Adán et al., 2018). Even though data journalism methods and a few monthly publications existed in Spain, the official introduction of data journalism in daily newspapers was first seen in 2013 in El Confidencial and after that El Español, El Mundo, El Periodico, and El Pais started to publish data-driven projects in the following years (Rivera Hernandez & Rojas Torrijos, 2016). Newsrooms began to recognize the value of data journalism in sports as new narratives emerged in the digital media. Articles with visualised data were helping readers understand, making sports simpler and becoming them familiar to the events (Rivera Hernandez & Rojas Torrijos, 2016). In Spain, the majority of newspapers now have separate data units. These units, however, collaborate with all sections in national newspapers. The adaptation cycle of sports data journalism is still ongoing within those circumstances. But it's seen that the majority of newspapers now publish online news that are more data and visual focused.

Until 2015, there were 21 digital sports news sources that paved the movement for sports data journalism. Olympics, women's football, tennis, basketball, football covered by these digital platforms (Manfredi Sánchez et al., 2015). AS Color, a digital edition for subscribers, preceded these measures, but with one notable difference. AS Color began as a printed newspaper in the early 1970s and, after nearly 40 years, transitioned to media with its digital edition. Sport joined the digital journalism community in 2014 when it released its digital edition, which was distributed on a monthly basis for tablets (Rojas Torrijos, 2015). Marca had a great deal of success. Marca Plus was launched in 2014 and has been one of Unidad Editorial's most successful emerging technology publications, "with 83 numbers in its first three years, an impressive positioning in the digital market: an average of 225,000 downloads per number and over fourteen million downloads since its launch." (Labarga-Adán et al., 2018). Multimedia content, data, gamification, and photojournalism were all used to create Marca Plus. In the field of data visualization and reader interaction with games, Marca Plus was achieving all data journalism practices. Following Marca, AS, and Sport, other newspapers, such as Mundo Deportivo El Periodico, Lavanguardia, and El Pais transformed their sports sections to more interactive and data driven.

5. Objectives and Method

This research aims to analyse data journalism practices in Spanish sports media and compare the adaptation process and preferable features in data journalism articles adapted by national and sports newspapers. In the light of these, this research aims to find answers to research objectives described below:

- *Research Objective 1:* To analyse technical features of sports data journalism articles in particular design principles, visualisation techniques, data sources, and trending topics.
- *Research Objective 2:* To identify the differences and similarities between the data journalism articles published by sports and national newspapers.

To find solid answers to proposed research objectives, we used content analysis of the published sports data journalism articles in the websites of Marca, Mundo Deportivo, Diario AS as sports newspapers and El Mundo, El Periodico, El Pais as national newspapers between 2017 and 2019.

This study aims to analyse the current level of sports data journalism in Spain with a quantitative approach. Quantitative approach led us to identify the technical features of sports data journalism articles including visualisation types, interactivity, data sources, trending topics, story properties and paved the way to achieve a comparative analysis between sports and national newspapers. We selected the time period as January 2017 to December 2019 based on the relatively late arrival of data journalism and its adaptation process in Spain and analysed sports

data journalism articles published in the websites of Marca, Mundo Deportivo, AS, El Periodico, El Mundo and El País. These newspapers were selected in order to their interest to publish data journalism articles, higher number of published articles compared to other newspapers and their daily circulation. Selected time period also presented the progress of data journalism in Spain and different paths followed by national and sports newspapers during this time. Finally, the final research sample was 1068 sports data-journalism based stories (Table 1).

Table 1. Research Sample.

Year	Newspaper						Total
	Marca	Mundo Deportivo	Diario AS	El Mundo	El Periodico	El Pais	
	n	n	n	n	n	n	n
2017	125	2	164	10	25	53	379
2018	70	29	184	39	24	87	433
2019	39	20	116	35	27	19	256
Total	234	51	464	84	76	159	1068

Source: Authors' elaboration from <https://www.marca.com>, <https://www.mundodeportivo.com>, <https://as.com>, <https://www.elmundo.es/deportes.html>, <https://elpais.com/deportes/>, <https://www.elperiodico.com/es/deportes/>

We created the codebook based on the relevant literature and the recent studies about data journalism (Appelgren, 2017; Loosen et al., 2017; Stalph, 2017; Tandoc Jr. & Oh, 2017; Veglis & Bratsas, 2017; Zhang & Feng, 2018) and we adapted the variables of related research to this study. The codebook consisted of 9 elements which are number of creators, data source, accessibility of data, topic, story property, number of visualisations, visualisation type, form of interactivity, and ratio of text and multimedia (Table 2).

Table 2. The Codebook.

Number of creators	0. Not mentioned; 1. 1 creator; 2. More than 1 creator
Data source	0. Not mentioned; 1. Mentioned
Accessibility of data	0. Not accessible; 1. Accessible
Topic	1. Football; 2. Basketball; 3. Tennis; 4. Motorsports; 5. Cycling; 6. Other
Story property	1. Comparison; 2. Connection and flow; 3. Changes over time; 4. Statistics; 5. Prediction; 6. Other
Number of visualisations	1. 1 visualisation; 2. 2 or 3 visualisations; 3. More than 3 visualisations
Visualisation type	1. Static; 2. Interactive

Form of interactivity	0. No interactive functions; 1. Transmissional interactivity (play, pause, zoom, buttons); 2. Consultational interactivity (timeline, menu, filters); 3. Conversational interactivity (personalization, inputs)
Ratio of text and multimedia	1. Mostly text; 2. Mostly multimedia; 3. Balanced; 4. Just multimedia

Source: Authors' elaboration based on The codebook is designed according to Appelgren, 2017; Loosen et al., 2017; Stalph, 2017; Tandoc Jr. & Oh, 2017; Veglis & Bratsas, 2017; Zhang & Feng, 2018.

Features of sports data journalism articles, changing trends in time, current data access level were identified through content analysis and sought solid answers to research objectives. Quantitative analysis of 1068 sports data journalism articles from selected newspapers presented the current adaptation level of data journalism in Spanish newsrooms and highlighted the differences between sports and national newspapers.

6. Results

Analysis of the articles published in selected newspapers presented the popular topics in sports data journalism and changing trends in years in both sports and national newspapers. Footballisation in sports journalism was observed in the analysed sample. Data journalism articles in football were seen as 65.2% in sports newspapers and 51.7% in national newspapers. Dominance of football articles was seen in each newspaper as 88.46% in Marca, 39.22% in Mundo Deportivo, 56.25% in Diario AS, 50% in El Mundo, 56.58% in El Periodico, 50.31% in El Pais. The analysis presented the visibility of other sports branches in sports and national newspapers. It was seen that basketball, tennis, motorsports, and cycling were published more by national newspapers while other sports including minor branches and Olympic sports were covered more by sports newspapers and sports newspapers were more football centric. (Table 3).

Table 3. Topics covered in sports data journalism articles between 2017-2019.

Topic	Sports newspapers		National newspapers	
	n	%	n	%
Football	488	65,2	165	51,7
Basketball	58	7,7	30	9,4
Tennis	39	5,2	50	15,7
Motorsports	54	7,2	34	10,7
Cycling	32	4,3	14	4,4
Other	78	10,4	26	8,1
Total	749	100	319	100

Source: Authors' elaboration from <https://www.marca.com>, <https://www.mundodeportivo.com>, <https://as.com>, <https://www.elmundo.es/deportes.html>, <https://elpais.com/deportes/>, <https://www.elperiodico.com/es/deportes/> based on Table 2.

Different narratives were seen in the sports data journalism articles published between 2017 and 2019. Comparison was the most preferred story property in both sports and national newspapers. Comparison based data journalism articles were mainly about significant games and key players. 34.7% of the data journalism articles in sports newspapers were based on comparison while 37.9% of the articles in national newspapers were comparison based. 'Other' as story property was seen as 20.2% in sports newspapers and 24.8% in national newspapers which includes photo-based explanations, mappings, listings, and drawings (Table 4). Statistics and change over time-based data journalism articles were seen at high rates as well and the least preferred story property was seen as prediction (Table 4). It is also notable that sports data journalism articles based on prediction were not seen in *Diario AS*, *El Mundo*, and *El País*.

Table 4. Story properties found in sports data journalism articles between 2017-2019.

Story property	Sports newspapers		National newspapers	
	n	%	n	%
Comparison	260	34,7	121	37,9
Connection & flow	49	6,5	12	3,8
Change over time	84	11,2	44	13,8
Statistics	198	26,4	61	19,1
Prediction	7	0,9	2	0,6
Other	151	20,2	79	24,8
Total	749	100	319	100

Source: Authors' elaboration from <https://www.marca.com>, <https://www.mundodeportivo.com>, <https://as.com>, <https://www.elmundo.es/deportes.html>, <https://elpais.com/deportes/>, <https://www.elperiodico.com/es/deportes/> based on Table 2.

Sports data journalism articles were created by journalists, data journalists and infographic artists based on the related departments and the number of article creators were variable. Both sports and national newspapers presented dominant rates in articles with 1 creator. 67% of the articles in sports newspapers were created by 1 journalist and/or infographic artist while the national newspapers were 51.7% (Table 5). It is also significant to highlight *El Mundo* in terms of mentioning the creators in each data journalism article which gives visibility to the author. Data journalism articles without any information of creators were not seen in *El Mundo*. Articles with more than 1 creator were mainly seen in more detailed stories and articles covered international events such as Olympics, Champions League, Euroleague and F1. These events were happening regularly, and the newspapers were planning their articles before the events because of the significance of events and audience interest.

Data sources are accepted as one of the most significant elements in data journalism articles. Based on the analysis of 6 newspapers, data journalism articles with data sources seen at high rates as 64.6% in sports newspapers and 74% in national newspapers (Table 6). In particular, *Mundo Deportivo* and *El Periódico* presented opposite results. Sports data journalism articles without data sources were seen with 98.04% in *Mundo Deportivo* and 97.37% in *El Periódico* in the analysed time period. *El Mundo* and *El País* were more likely to share data sources in articles with 97.62% in *El Mundo* and 95.6% in *El País*. It's worth noting that accessibility of data sources

was not seen in Mundo Deportivo, Diario AS, El Mundo, El Periodico and El País however Marca was the only newspaper that gives access to data sources in 4.7% of the articles.

Table 5. Number of creators in sports data journalism articles between 2017-2019.

	Sports newspapers		National newspapers	
	n	%	n	%
Not mentioned	94	12,6	28	8,8
1 creator	502	67,0	165	51,7
More than 1 creator	153	20,4	126	39,5
Total	749	100	319	100

Source: Authors' elaboration from <https://www.marca.com>, <https://www.mundodeportivo.com>, <https://as.com>, <https://www.elmundo.es/deportes.html>, <https://elpais.com/deportes/>, <https://www.elperiodico.com/es/deportes/> based on Table 2.

Table 6. Data sources of sports data journalism articles between 2017-2019.

	Sports newspapers		National newspapers	
	n	%	n	%
Not mentioned	265	35,4	83	26,0
Mentioned	484	64,6	236	74,0
Total	749	100	319	100

Source: Authors' elaboration from <https://www.marca.com>, <https://www.mundodeportivo.com>, <https://as.com>, <https://www.elmundo.es/deportes.html>, <https://elpais.com/deportes/>, <https://www.elperiodico.com/es/deportes/> based on Table 2.

A diversity in the number of visualisations used in articles were seen in sports newspapers while national newspapers were publishing data journalism articles with 1 visualisation dominantly (Table 7). In addition, the number of visualisations presented preferred styles by newspapers. Sports data journalism articles published by Diario AS showed balanced rates in number of visualisations while El Mundo and El Periodico were preferring articles with 1 visualisation. Also, articles with more than 3 visualisations were not preferred by these newspapers and seen with 5.26% in El Periodico and 0 in El Mundo. Use of more than 3 visualisations were mostly seen in the articles with transmissional and consultational interactivity in the topics of big events, end of season comparisons, and league statistics. Visualisations in sports data journalism articles in the analysed period were mostly static and did not allow readers to play and interact with visualisations. It was seen that 84% of the visualisations in sports newspapers and 96.6% of the visualisations in national newspapers were static (Table 8).

Table 7. Number of visualisations in sports data journalism articles between 2017-2019.

	Sports newspapers		National newspapers	
	n	%	n	%
Number of visualisations				
1 visualisation	273	36,5	160	50,1
2-3 visualisations	218	29,1	88	27,6
More than 3 visualisations	258	34,44	71	22,3
Total	749	100	319	100

Source: Authors' elaboration from <https://www.marca.com>, <https://www.mundodeportivo.com>, <https://as.com>, <https://www.elmundo.es/deportes.html>, <https://elpais.com/deportes/>, <https://www.elperiodico.com/es/deportes/> based on Table 2.

Table 8. Visualisation types in sports data journalism articles between 2017-2019.

Visualisation type	Sports newspapers		National newspapers	
	n	%	n	%
Static	629	84,0	308	96,6
Interactive	120	16,0	11	3,4
Total	749	100	319	100

Source: Authors' elaboration from <https://www.marca.com>, <https://www.mundodeportivo.com>, <https://as.com>, <https://www.elmundo.es/deportes.html>, <https://elpais.com/deportes/>, <https://www.elperiodico.com/es/deportes/> based on Table 2.

Form of interactivity in sports data journalism articles presented a difference in sports and national newspapers. Articles with no interactive functions were dominant with 80.2% in sports and 94.6% in national newspapers but the use of transmissional, consultational, and conversational interactivity presented different rates (Table 9). Sports newspapers were more interested in publishing articles with transmissional interactivity which includes playing, pausing, zooming and button options than national newspapers. However, Diario AS was seen as the primary newspaper publishing articles with no interactive functions. 99.78% of the articles published by Diario AS had no interactive functions. In addition, the analysis presented the lack of consultational and conversational interactivity in national newspapers. 1.6% of the articles published by national newspapers had consultational interactivity and all these articles were seen in El Pais. Conversational interactivity was not observed in national newspapers.

Ratio of text and multimedia in the analysed sports data journalism articles presented opposite results. Articles based on mostly multimedia was seen dominantly in sports newspapers and followed by articles with balanced use of text and multimedia (Table 10). Sports newspapers were interested in storytelling mainly with visualisations and less text. However, there was no balance in national newspapers in terms of ratio of text and multimedia. National newspapers were publishing data journalism articles either with just multimedia or mostly text.

Table 9. Form of interactivity seen in sports data journalism articles between 2017-2019.

Form of interactivity	Sports newspapers		National newspapers	
	n	%	n	%
No interactive functions	601	80,2	302	94,6
Transmissional	81	10,8	12	3,8
Consultational	61	8,1	5	1,6
Conversational	6	0,8	0	0
Total	749	100	319	100

Source: Authors' elaboration from <https://www.marca.com>, <https://www.mundodeportivo.com>, <https://as.com>, <https://www.elmundo.es/deportes.html>, <https://elpais.com/deportes/>, <https://www.elperiodico.com/es/deportes/> based on Table 2.

Table 10. Ratio of text and multimedia in sports data journalism articles between 2017-2019.

Ratio of text and multimedia	Sports newspapers		National newspapers	
	n	%	n	%
Mostly text	41	5,5	100	31,4
Mostly multimedia	434	57,9	32	10,0
Balanced	202	27,0	55	17,2
Just multimedia	72	9,6	132	41,4
Total	749	100	319	100

Source: Authors' elaboration from <https://www.marca.com>, <https://www.mundodeportivo.com>, <https://as.com>, <https://www.elmundo.es/deportes.html>, <https://elpais.com/deportes/>, <https://www.elperiodico.com/es/deportes/> based on Table 2.

6. Discussion and Conclusion

The analysis presented the features of sports data journalism articles in selected newspapers and a comparative perspective. Based on the research, similarities, and differences in technical characteristics of sports data journalism articles published by sports and national newspapers were observed. These results were significant to identify the current adaptation level of data journalism in Spanish newsrooms. Results presented an ongoing progress in Spanish newsrooms and variable perspectives of newspapers towards sports data journalism.

We identified the footballisation in Spanish sports media through the number of data journalism articles about football during the selected time period and this dominance was based on the amount of audience and reader interest. Cycling was the least visible sports branch in Spanish data journalism. We observed that all newspapers were interested in publishing football but in time national newspapers started to publish more articles in other topics too. The analysis presented increased numbers of tennis and basketball articles in El Mundo, tennis, and motorsports articles in El Periodico and men's tennis articles in El Pais in the analysed time

period. We discovered that El Mundo's data journalism stories were mostly focused on single events such as the FIFA World Cup and the Olympics. Increased percentages of certain topics over time were found to be due to global events, according to the study. Football and "other" sports had the most significant rises.

We analysed that all newspapers were publishing comparison based articles regularly. Comparison based data journalism articles were seen at higher rates compared to other story properties. However, other story properties presented a variety in results. Sports newspapers were publishing comparison, statistics, and 'other' dominantly but connection and flow and change over time based articles were also preferred. Dominance of comparison, statistics and 'other' based data journalism articles were observed in national newspapers but the interests towards connection and flow was lower and connection and flow was higher than sports newspapers. Use of other techniques was seen in El Periodico and El Pais intensively. Sports were offering advanced statistics however not all readers were qualified enough to understand these statistics. Sports were more comprehensible to readers through data journalism articles and based on this idea we identified that Spanish newspapers were choosing comparison and statistics-based articles to make it simpler. We observed a lack of education in newsrooms based on the dominance of articles based on 'other' techniques such as photo-based explanations, mappings, drawings because advanced knowledge in data journalism tools were not needed to create these types of articles.

Sports and national newspapers presented similar results in terms of number of creators. Data journalism articles were written and visualised by 1 creator in both sports and national newspapers dominantly. The articles with more than 1 creator were observed at higher rates in national newspapers. Most of the articles were written by a single creator but there was a connection between the number of creators and the number of visualisations. While the number of visualisations were increasing, the numbers of creators were increasing at the same time. It was seen that El Pais increased the number of visualisations and included more creators in the articles during the analysed period. It was seen that data journalism articles in national newspapers had more information about data sources than sports newspapers. These data sources were generally private data providers, transparency portal, and open data. When there were few data sources, national newspapers were mentioning all data sources used in the article. Sports newspapers were generally publishing the name of the single data source which is an open portal or a data came from a provider. When there were few data sources sports newspapers were mentioning the data source as "own elaboration." We observed that the data sources of newspapers were not accessible except a few articles in Marca.

Sports newspapers were more interested in publishing more visualisations than national newspapers. Data journalism articles with 1 visualisation and longer texts were seen in the analysis of national newspapers while sports newspapers were interested in publishing higher numbers of visualisations based on less text. Sports newspapers were aiming to tell the story mainly with visualisations and supportive text while national newspapers were using less visualisations and dominant amounts of text. We also identified the difference in use of interactivity between sports and national newspapers. Both sports and national newspapers were mainly publishing data journalism articles with no interactive functions however sports newspapers were trying different techniques to communicate with readers. Sports newspapers were trying to use all types of interactivity which are transmissional, consultational, and conversational; national newspapers were more depending on classical formats. Lower numbers of articles with transmissional and consultational interactivity were seen in the analysis of the articles in national newspapers while there was not an article with conversational interactivity.

Sports is offering a variety of stories behind advanced statistics, but we identified that the progress of sports data journalism in Spanish newsrooms was not at the desired level compared to other countries. Possible reasons for this slow progress can be discussed in order to identify the current level in a better way. Daily workload, lack of education and traditional journalistic

environment in Spain can be assumed as the reasons for the progress in recent years. Data journalism tools and new narrative techniques are rapidly growing, and it can be challenging to follow these new trends for Spanish newspapers. Also, daily workload is a significant point to create data journalism articles to spend the necessary time for planning, research, and visualising. At this point, we observed that the adaptation of data journalism in sports newspapers was at a higher level than national newspapers. These observations can be discussed under publishing in specialized topics. National newspapers are publishing articles in several topics and the newsrooms need to organize the workload for data journalism articles for all topics. However, sports newspapers are publishing almost the same number of articles but specialized just in sports. Data journalists, infographic artists, and reporters are specialized in sports and the daily workload is not affected by Spain's daily agenda like national newspapers.

Consequently, the current level of sports data journalism in Spain is promising for the future years because of the global interest to Spanish sports industry, available data, and players' and teams' popularity around the world. However, newspapers should create separate data journalism departments in order to expedite this progress and more investment is needed for data journalism education.

References

- Appelgren, E., & Nygren, G. (2014). Data Journalism in Sweden: Introducing new methods and genres of journalism into old organizations. *Digital Journalism*, 2(3), 394-405. <https://doi.org/10.1080/21670811.2014.884344>
- Appelgren, E. (2017). An Illusion of Interactivity: The paternalistic side of data journalism. *Journalism Practice*, 12(3), 308-325. <https://doi.org/10.1080/17512786.2017.1299032>
- Appelgren, E., & Salaverria, R. (2018). The Promise of the Transparency Culture. *Journalism Practice*, 12(8), 986-996. <https://doi.org/10.1080/17512786.2018.1511823>
- Arias, D., Sánchez García, P., & Redondo, M. (2018). Profile of the data journalist in Spain: training, sources and tools. *Doxa Comunicación*, 27, 273-293.
- Baack, S. (2011). A new style of news reporting: Wikileaks and data-driven journalism. *Cyborg Subjects*, 10, 1-10.
- Blanco Castilla, E., & Quesada, M. (Eds.) (2016). *Periodismo de datos*. Cuadernos Artesanos de Comunicación. Latina.
- Borges Rey, E. (2016). Unravelling Data Journalism. *Journalism Practice*, 10(7), 833-843. <https://doi.org/10.1080/17512786.2016.1159921>
- Charbonneaux, J., & Gkouskou-Giannakou, P. (2015). "Data Journalism", an investigation Practice?: A glance at the German and Greek cases. *Brazilian Journalism Research*, 11(2), 244-267. <https://doi.org/10.25200/BJR.v11n2.2015.855>
- Coddington, M. (2015). Clarifying Journalism's Quantitative Turn. *Digital Journalism*, 3(3), 331-348. <https://doi.org/10.1080/21670811.2014.976400>
- Craig, D., Zion, L. (2014). *Ethics for Digital Journalists: Emerging Best Practices*. Routledge.

- De Maeyer, J., Libert, M., Domingo, D., Heinderyckx, F., & Le Cam, F. (2015). Waiting for Data Journalism: A qualitative assessment of the anecdotal take-up of data journalism in French-speaking Belgium. *Digital Journalism*, 3(3), 432-446. <https://doi.org/10.1080/21670811.2014.976415>
- Ferreras Rodriguez, E.M. (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
- Garrison, B. (1998). *Computer-Assisted Reporting*. Lawrence Erlbaum Associates.
- Grassler, M. (2018). *El rol del periodista de datos en el proceso de los sistemas de gestión y decisión pública y en la recuperación de la confianza entre el ciudadano y las instituciones públicas*. Universidad Autónoma de Barcelona. <http://hdl.handle.net/10803/462769>
- Hammond, P. (2015). From computer-Assisted to data-Driven: Journalism and Big Data. *Journalism: Theory, Practice & Criticism*, 18(4), 408-424. <https://doi.org/10.1177/1464884915620205>
- Howard, A. (2014). *The Art and Science of Data-Driven Journalism*. Columbia Journalism School.
- Labarga Adan, I., Gonzalez Diez, L., & Perez Cuadrado, P. (2018). Digital native magazines in the field of Sports in Spain: the case of MARCA Plus. *Revista Mediterránea de Comunicación*, 9(1), 451-463.
- La-Rosa, L., & Sandoval-Martín, T. (2016). The Transparency Law's insufficiency for Data Journalism's practices in Spain. *Revista Latina de Comunicación Social*, 71, 1208-1229.
- Loosen, W., Reimer, J., & De Silva-Schmidt, F. (2017). Data-Driven Reporting – an On-Going (R) Evolution? A Longitudinal Analysis of Projects Nominated for the Data Journalism Awards 2013– 2015. *Journal of Hans-Bredow-Institut for Media Research*, 41. <https://doi.org/10.21241/ssoar.71776>
- Manfredi Sánchez, J.L., Rojas Torrijos, J.L., & Herranz de la Casa, J.M. (2015). Entrepreneurial journalism: Sports journalism in Spain. *Revista Latina de Comunicación Social*, 70, 69-90.
- Meyer, P. (2002). *Precision Journalism: A Reporter's Introduction to Social Science Methods*. Rowman & Littlefield.
- Michalski, D. (2016). *Reader Engagement with Data Journalism: Comparing the Guardian and Washington Post's Coverage of People Killed by Police*. University of Nevada. <http://dx.doi.org/10.34917/9302952>
- Olivo, M.T. (2015). *Visualizing the News: An Analysis of a Year in Interactive News from The New York Times and The Washington Post*. Rochester Institute of Technology. <https://scholarworks.rit.edu/theses/8910>
- Porlezza, C., & Splendore, S. (2019). From Open Journalism to Closed Data: Data Journalism in Italy. *Digital Journalism*, 7(9), 1230-1252. <https://doi.org/10.1080/21670811.2019.1657778>
- Rivera Hernandez, A., & Rojas Torrijos, J.L. (2016). El Español and El Confidencial, models of data sports journalism in Spanish native digital news media. *Doxa Comunicación*, 23, 1-23.
- Rojas Torrijos, J.L. (2015). New perspectives on Sports Journalism in Spain: first magazines for tablets and smartphones in daily newspapers Marca and Sport. *Fonseca Journal of Communication*, 10(10), 29-49.
- Rojas Torrijos, J.L. (2016). El Español and El Confidencial, models of data sports journalism in Spanish native digital news media. *Doxa Comunicación*, 23, 1-23.

Schulze, T. (2015). *Data Journalism, Millennials & Social Networks*. Universidade Fernando Pessoa. <http://hdl.handle.net/10284/5039>

Stalph, F. (2017). 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 Jr, E.C., & Oh, S.K. (2017). 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>

Veglis, A., & Bratsas, C. (2017). Towards a Taxonomy of Data Journalism. *Journal of Media Critiques*, 3(11), 109-121.

Zhang, S., & Feng, J. (2018). 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>



© Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0)
<https://creativecommons.org/licenses/by-nc-nd/4.0/>