

# Professional domains of data journalism within the digital media ecosystem: an approach to the Spanish context and to leading media in the English-speaking world

*Ámbitos profesionales del periodismo de datos dentro del ecosistema mediático digital: una aproximación al contexto español y a los medios de referencia en el mundo anglosajón*

ISSN: 1696-019X / e-ISSN: 2386-3978



**Marta Saavedra Llamas.** PhD accredited with a six-year period of active research (2015-2020). She is Vice Dean of the Faculty of Communication and Arts of the Antonio de Nebrija University since March 2021. She is part of the Innomedia Research Group. Previously, she was part of the privately funded Journalistic Analysis and Anticipation Research Group, focused on health and environmental journalism. Her main lines of research are the communication industries, communication strategies, the social audience and professional profiles in communication. She has published four books in leading publishers, fourteen articles in indexed journals, and thirty book chapters. She has participated in numerous international conferences.  
Nebrija University, Spain  
msaavedr@nebrija.es  
ORCID: 0000-0002-7762-9996



**Mercedes Herrero de la Fuente.** PhD in Information Sciences (UCM). Member of the Innomedia research group and member of the R+D+I project “Disability and digital skills in the audiovisual sector”. She has been a research fellow at: Cornell University (USA), University of Salford (United Kingdom), Radboud Universiteit (Netherlands) and Karlova Univerzita (Czech Republic). She has published articles in highly indexed journals and participated in numerous international conferences. Her lines of research include the application of new technologies to informative discourse, disinformation in social networks, and the new professional profiles. Professor accredited for public and private universities, she teaches classes at the Antonio de Nebrija University, in undergraduate and postgraduate degrees.  
Nebrija University, Spain  
mherrero@nebrija.es  
ORCID: 0000-0002-5361-9056



**Eduardo Castillo Lozano.** Coordinator of the Journalism degree at the Antonio de Nebrija University. Degree in Journalism from the San Pablo CEU University. Doctoral student in Communication at the UCM, where he is conducting research on data journalism and new professional profiles. As a professor at the Nebrija University, he teaches subjects related to journalistic writing, communication skills in writing and public speaking, or locution. Radio journalist specializing in economics. He is currently the director and presenter of “Afterwork” on Capital Radio (a program for which he received an Antena de Oro in 2017). He has been the presenter of the TVE La2 television program “Conecta con el Mercado”.  
Nebrija University, Spain  
ecastill@nebrija.es  
ORCID: 0000-0003-1726-2712

## How to cite this article:

Saavedra Llamas, M.; Herrero de la Fuente, M. and Castillo Lozano, E. (2023). Professional domains of data journalism within the digital media ecosystem: an approach to the Spanish context and to leading media in the English-speaking world. *Doxa Comunicación*, 37, pp. 341-358.

<https://doi.org/10.31921/doxacom.n37a1853>



This content is published under Creative Commons Attribution Non-Commercial License. International License

**Abstract:**

Technological transformation is driving the evolution of the media ecosystem, in which data journalism is consolidating as an area of development. Among the aspects associated to this journalistic modality, data visualization, news verification and investigative journalism stand out. In this context, newsrooms made up of interdisciplinary teams are flourishing, where technological skills are gaining ground. Using the focus group technique, we collected the qualified opinions of twelve journalists from Spanish media, whose daily work is related to data journalism. We completed this with two in-depth interviews with journalism professionals from The Washington Post and The Guardian, both pioneering media in its implementation. Our results indicate that the most valued skills are center on database exploitation, big data analysis, visualization, and fact-checking. In Spain, data journalism is not advancing at the desired speed, because it is difficult to find qualified professionals and investments in technology is insufficient. In the Anglo-Saxon world, progress is remarkable, with a significant rise in the Graphics department. In both scenarios, data-related skills improve journalists' employability and career advancement.

**Keywords:**

Data journalism; data visualization; fact-checking; professional profiles; professional skills.

**Resumen:**

*La transformación tecnológica impulsa una evolución del ecosistema mediático, en la que el periodismo de datos se consolida como área de desarrollo. Entre las vertientes ligadas a esta modalidad periodística destacan la visualización de datos, la verificación informativa y el periodismo de investigación. En este contexto florecen las redacciones integradas por equipos interdisciplinarios, donde las habilidades tecnológicas ganan terreno. Con la técnica del focus group recabamos la opinión cualificada de doce periodistas de medios españoles, cuyo trabajo diario está relacionado con el periodismo de datos. Completamos lo anterior con dos entrevistas en profundidad a profesionales de The Washington Post y The Guardian, medios relevantes en su implantación. Nuestros resultados indican que las competencias más valoradas se centran en la explotación de bases de datos, el análisis de big data, la visualización y el fact-checking. En España, el avance del periodismo de datos no se produce a la velocidad deseada, porque resulta difícil encontrar profesionales cualificados y la inversión en tecnología es insuficiente. En el mundo anglosajón el progreso es notable, con especial auge del departamento de Graphics. En ambos escenarios, las competencias relacionadas con datos mejoran la empleabilidad de los periodistas y su promoción profesional.*

**Palabras clave:**

*Periodismo de datos; visualización de datos; verificación; perfiles profesionales; competencias profesionales.*

**1. Introduction**

The impact of technological changes on the media ecosystem has become more severe since the year 2020. The pandemic experienced worldwide accelerated the ongoing processes and confirmed trends that are now true, and which must be dealt with in daily journalistic work. Many national and international reports, such as those by *Reuters Institute for the Study of Journalism* or the Observatorio Nacional de Tecnología y Sociedad (ONTS, National Observatory of Technology and Society), have detected an evolution towards hybrid newsrooms<sup>1</sup>, a change in the journalistic routines, and a growing role of data in the process of the creation of information and its posterior verification, in a moment in time in which disinformation has risen as a threat against the future of media (Amoedo-Casais et al., 2022; Newman, 2022; Cherubini et al., 2021; Velasco-Jones et al., 2021). In the present section, we will discuss all of these matters, being aware that we are currently living in a moment in

1 Here, as in section 1.1 the term hybrid is used in the sense of teams in which part of the professionals work in person in the newsroom, while the other part works remotely.

time in which the concept of liquid modernity (Bauman, 2005) seems to have overcome what Scolari (2021) named “gaseous metaphor”, which describes a media ecology in which we find the multiplication of actors, formats, and practices. This is a complex and shifting framework, in which transformations are fast and very often global.

### 1.1. *Data in journalistic practice: data journalism*

Data journalism comes from the demand for transparency of public administrations promoted by the movement *open data* and the great effect of *big data* on every aspect of society. At first, it became important in specific media outlets in the United States and the United Kingdom (Fink and Anderson, 2014; Hewett, 2013; Ferreras-Rodríguez, 2013), afterwards extending to other countries, and it is now a strong trend in conventional media in all of Europe, as we verified during the health crisis in 2020 (Córdoba-Cabús et al., 2020), and in digital natives (Vizoso-García, 2021).

The use of data seeks to find interesting and quality content, and involves the use of tools for its extraction and management, as well as its posterior visualization. This transforms the manner in which to obtain information, and also the writing and presentation of the contents (Ferreras-Rodríguez, 2013; Saavedra-Llamas et al., 2019).

Data journalism is considered the continuation of precision journalism, among other associated specializations. Chaparro-Domínguez (2013) and Ferreras-Rodríguez (2013) so argue, while Dader-García and Gómez-Fernández (1993) point out that just as in-depth journalism, precision journalism is derived from investigative journalism.

Philip Meyer is one of the main leaders of this journalistic evolution, as in the 60s, he began to develop journalism stories based on the analysis and treatment of data. The particular nature of this type of journalism was that the scientific methodology from the Social Sciences was applied to the exercise of journalism in an innovative manner. Meyer defined this specialty as:

The application of social and behavioral science research methods to the practice of journalism [...] The tools of sampling, computer analysis, and statistical inference increased the traditional power of the reporter without changing the nature of his or her mission-to find the facts, to understand them, and to explain them without wasting time. (2002: 2-3).

Meyer proposes a work scheme that would be valid today in the processes of creation of stories based on data journalism. First, one must know how to collect data, store them, be able to retrieve them, analyze them, reduce them, and summarize them, and most importantly, know how to tell a story with them.

Likewise, Dader-García (1997) describes a series of characteristics for precision journalism, that are still current for data journalism, such as that numerical quantification, scientifically verified, be the central element of the journalistic story.

Although *The Washington Post* or *The Guardian* are now leaders in the area of data journalism, and have large departments, these media outlets have had numerous forerunners in the development of data-driven stories. This is the case of the works published in the newspaper *Detroit Free Press* in 1967 (López-López, 2016), or those by *Providence Journal* in 1984 (Casal-Cortizas, 2010).

The technological revolution has been a key aspect for the development of this specialization. In fact, precision journalism was superseded by computer-assisted reporting (CAR). Gehrke and Mielniczuk (2017) also pointed to Meyer as the creator of CAR, although they argued that the use of a computer was not the key in precision journalism, but instead it was the use of the

scientific method, above technology. On the other hand, DeFleur (1997) indicates that CAR is a precursor of data journalism, and considers it as a discipline that performs computer analyses starting from records from the Government and other public agencies, systematically.

Dader-García and Gómez-Fernández (1993) point out that the evolution of computer systems promoted the evolution from CAR towards the so-called *Database Journalism*, a specialty that already includes in its definition the concept of data, and bases its methodology in the screening of databases supported by powerful computer systems, which allow the exploration and comparison of numerical and non-numerical content.

However, the authentic technological revolution that grants meaning to data journalism has been the emergence of the era of *big data* and the digitalization of processes. According to La Rosa-Barrolleta (2019), data journalism differentiates itself from precision journalism and CAR, in that it incorporates problems and techniques that have emerged from the era of *big data*.

Related with this, the new specialized competences and skills demanded by the media are centered in areas such as new technologies, statistics, the management of databases, the screening techniques and analysis, the visualization of the information, and multimedia narratives (López-García et al., 2016a).

The studies on the new professional profiles that have emerged in the area of journalism are numerous and interesting, among which we find those by López-García (2001), Vivar-Zurita (2011) and Sánchez-García et al. (2015). In this specific case, the work by Ferreras-Rodríguez (2012) deserves a special mention, as it is one of the first that referred to the data journalist as a new professional alternative with good job prospects. These and other studies introduce the idea of the multi-purpose and versatile profile (Salaverría-Aliaga and Martínez-Costa, 2021; Gómez-Calderón et al., 2017; Heredero-Díaz and Reyes-Sánchez, 2016) with cross-sectional training, in which knowledge on statistics and programming are highly valued (Saavedra-Llamas and Toledano Cuervas-Mons, 2022). These professionals work in multi-disciplinary teams, with a marked technological component (Saavedra-Llamas et al., 2020), where “the traditional division between engineers and journalists doesn’t make sense” (Manfredi-Sánchez et al., 2019: 1643). These teams are the result of different processes of technological innovation adopted from newsrooms (García-Avilés et al., 2019; Nee, 2013), some of which the fruit of the development of data journalism. González-Alba (2022) includes an interview given to the managing editor of the *Washington Post*, Kat Downs Mulder, who explained that this newspaper had more than two hundred *software engineers* employed in different areas, who work together with other journalists, looking for the best product and the ideal way to present it to the audience. As a postscript, we must mention the term “adaptive journalist,” coined by Palau-Sampio in 2016 and mentioned by García-Avilés, who designates these professionals as “open to experimentation in multidisciplinary teams in which it is common to find collaboration between technical and journalist profiles” (2021: 10).

The first data journalism initiatives in Spain surged from this idea of multidisciplinary teams. In 2001, at the Medialab Prado space, a cultural and social innovation center, the first workshops on data journalism took place, led by leaders such as Mar Cabra, Alberto Cairo or David Cabo (Ferreras-Rodríguez, 2016). The latter would launch the CIVIO Foundation in 2012, a non-governmental entity destined to offering information on transparency based on data journalism, and which employs many journalists, programmers, and designers.

The supply of individuals talented on the use technology and data are limited, and are demanded by many sectors, with which media outlets compete. Cherubini et al. (2021) reveals that almost half of the media outlets consulted have difficulties in hiring professionals with the necessary qualifications. The main problems are centered on the more specific positions, “such as software engineers or data scientists” (Newman, 2022:22). Some companies have proposed work flexibility as incentives, highlighting remote working for performing of tasks as one of the most recurring elements to attract talent in hybrid newsrooms (Cherubini et al., 2021). Perhaps due to this, and according to the same study, 34% of those consulted will maintain a hybrid model, while 57% have not made a decision on this aspect (2021: 6). Engebretsen et al. (2018) affirmed, in their research on the practice of data journalism in the Nordic regions, that “priorities are given to the specialists in this field when hiring” (2018: 13).

Within the specific area of data journalism, data visualization and verification appear as areas that are experiencing the greatest growth, as shown by the news coverage on COVID-19. This area also provides new possibilities for investigative journalism, a very important facet in the history of specific media.

The presentation of data through infographics focused much of the information during the pandemic, gaining ground as compared to other types of images, especially photography. These visual resources often appear as independent information elements, which grants them with special relevance, state Salvat-Martinrey and Aranda-Martínez (2020) in their work centered on the main Spanish newspapers during the coronavirus crisis. López-García et al. (2016) highlight the added value provided by data analysis techniques and their visualization, to news stories. Engebretsen et al. (2018) specifically pointed out *mobile first* as one of the dominant criteria. This is translated into a simplification of the infographics, with less possibilities of interaction, as they seek to be viable and operational for their consultation through *smartphones*, to multiply their viewing (Engebretsen et al., 2018).

As for verification, it is also revealed as an indispensable tasks in a context dominated by disinformation. Cerezo (2021) describes it as among the main lines through which journalism currently advances, to meet a growing demand for truthful and quality information through different formats and channels. In 2005, Channel 4 News inaugurated, in the United Kingdom, the first verification blog in Europe. Its example was followed in the entire continent, and by 2010, there were already active verification platforms in 10 European countries (Rodríguez-Martínez et al., 2021: 29). Presently, verification journalism is being developed through conventional media, independent agencies, and other entities, counting with international organizations that define shared guidelines to perform this work in a manner that is rigorous and recognized, such as the *International Fact-Checking Network* (IFCN) or *The Trust Project* (Vázquez-Medel and Pulido-Rodríguez, 2021). The verification work emerges as a response from the profession against the extended rumors in the digital context (López-García et al., 2016), which affects transcendental matters, such as the continuity of the democratic system (Herrero-De-La Fuente and Ríos-Calvo, 2022). On the other hand, disinformation is one of the main concerns of citizens, with 67% of users declaring their concerns on this matter in Spain. This number is higher than the mean of 58% recorded in 46 countries, as shown in the last *Digital News Report Spain 2021* (Amoedo et al., 2022: 6).

The application of data to investigative journalism provides new possibilities for the publication of content from information obtained through data searches and filtering techniques, which could become important news articles for the audience,

and in general, for the interest of the population (Picard, 2017; Martinisi 2013; Flores-Vivar and Salinas-Aguilar, 2013). This journalistic branch is connected, as the previous one, with the need to manage truthful information that provides clarity in a moment in time characterized by media noise and growing polarization.

In Spain, there has been a notable development of data journalism in the main communication media, such as *El Confidencial*, *Eldiario.es* or *El Español*, in which it is frequent to find news articles that offer information, graphics, and interactions on current events (Rubio-Jordán, 2021).

## 2. Objectives and methodology

The digitalization of communication media, as we have observed in the introduction, has provoked an unprecedented transformation in information and entertainment industries. Changes have been made in the media ecosystem, the organization of work teams, journalism routines, and professional skills. One of the positive aspects is the proliferation of different professional profiles that must tend to the new dimensions of journalism, and which were born, specifically, from the new possibilities provided by technology. Thus, within the digital area, we observed the strong surge of a concept of data journalism closely associated with the verification of information, exploitation of *big data* for informational aims, and the transformation of data into stories, together with its attractive presentation for audiences.

The main objective of the present article is to delve into this journalistic specialization, and to discover if, along with it, new professional profiles have proliferated, what functions they play, and what skills are needed. Likewise, the intention is to observe if this renewal of profiles has resulted in the re-organization of newsrooms, and if it has resulted in the hybridization of work teams.

For this, the study was structured into three phases. The first was descriptive, and based on a bibliographical and documental exploratory study to be find the context within which we could frame our object of study. The second and third phases were qualitative, and complementary to each other. Firstly, the *focus group* technique was utilized to address, together with specialized professionals, the study of the new profiles.

The *focus group* technique is very valuable in the social sciences, as it allows receiving feedback about diverse subjects, as stated by Kotler and Armstrong (2016). Thus, we believe that it will be highly useful for the present work, as we can compile professional impressions and experiences of journalists from different media outlets, and verify if there is consensus, or if the data-associated profiles are more or less developed according to the sector. Also, as Garvin (2008) explains, this qualitative tool offers a greater field of information than quantitative techniques, as those who are interviewed not only answer close-ended questions, but their interaction can generate new connections. On their part, Edmunds (1999) argues that the great advantage of the *focus group* is that it allows us to discover what meanings are attributed by different individuals to specific subjects, ideas, or concepts.

Valles-Martínez (2002) reminds us to select the participants well, considering a series of variables that justify their selection. Thus, in the present study, the main concurrent selection variable was that the participants had to be active in the area of data journalism, on order to be able to explain their functions, determine their fundamental skills, and describe their different roles.

A secondary concurrent variable was added, in that the experts had to have a trajectory longer than three years in this area, to be able to observe the evolution of the data journalism specialization. The representation of men and women was sought, as well as the incorporation of different communication industries. With respect to this, different categories were established: Communication media with a digital section (A); Communication media with a data journalism section (B); Special data units created within the corporation or medium (C), and Verification agencies and specialized entities (D).

Below, we present the experts selected, who were given a code to simplify the *verbatim* texts extracted.

**Table 1. Experts who participated in the *focus group***

Expert and profile	Type	Medium	Code
Llerena, María José; digital development manager	A	<i>Onda Cero</i>	E1
Pomar, Julio; head of digital analytics		<i>ABC</i>	E2
Ramos, David; editor-in-chief of content and digital strategy		<i>Cadena Ser</i>	E3
Castellón, Joaquín; coordinator of research and data	B	<i>La Sexta</i>	E4
Escudero, Jesús; data journalist specialized in statistics		<i>El Confidencial</i>	E5
Gallego, Francisco; <i>Head of digital analytics y data management</i>		<i>Unidad Editorial</i>	E6
Escalona, Patricia; <i>fact-checker</i> in <i>Verifica Antena 3</i>	C	<i>Antena 3</i>	E7
García, Desirée; manager of <i>Efe Verifica</i>		<i>EFE</i>	E8
Guisado, Paula; head of <i>Datos RTVE</i>		<i>RTVE</i>	E9
Calle, Ignacio; director <sup>2</sup>	D	<i>Maldito Dato</i>	E10
Delgado, Antonio; co-founder		<i>Datadista</i>	E11
Martín, Alba; journalist specialized in visualization of data		<i>Newtral</i>	E12

**Source:** created by the authors from the *curriculum vitae* of the experts

The third phase utilized the in-depth interview technique, as we consider it valuable for obtaining the testimony of data journalists who perform their activity outside of Spain, in markets where the specialization is more developed, and even in media outlets that are considered leaders worldwide. Thus, the study also counted with the participation from the United States and the United Kingdom, and the prestigious media outlets *The Washington Post* and *The Guardian*, respectively.

<sup>2</sup> This position was current when the *focus group* took place, but as of August, 2022, he is the research director at *Público.es*.

It must be mentioned that these two countries have been the pioneers in the development of data journalism, or, as argued by Fink and Anderson (2015), in the development of computational, innovative, and revolutionary journalism work. Also, *The Washington Post* is considered a worldwide example of the technological development of communication media and funding for journalism (Zapater-Delgado, 2021). On its part, *The Guardian* has developed the specialty object of study with the creation of *Datablog*, a channel considered by many theorists as fundamental in the evolution of precision journalism to data journalism Chaparro-Domínguez, 2013). Also, as pointed out by López-García et al. (2016a), it is considered a benchmark in the English-speaking world in working with data and its visualization. Likewise, both publishers are underlined by López-García et al. (2016b) as benchmarks in the processes of verification, through its analysis of their *fact checking* models.

As for the method utilized in the third phase, Peña (2013) explains that an in-depth interview is a scientific-professional conversation through which we can obtain information that will help us analyze a social context. The author determines, in agreement with Kvale (2011) and Denzin and Lincoln (2000), that it is the most convenient technique within qualitative models.

This is not a spontaneous dialogue, as Valles-Martínez (2002) explains, but it is founded on a structured process in which the receiver of information –the researchers or interviewer– must select the person questioned –the expert or interviewee– based on a set of variables and a solid argument. Likewise, a reflection must be made on the subjects that are dealt with, and create a space of exchange that facilitates the interaction between both agents.

Given these indications, the interviewees were:

Table 2. Experts interviewed

In-depth interview 1	In-depth interview 2
Esteban, Chiqui Graphics Director <i>The Washington Post</i> USA	Gutiérrez, Pablo <i>Visual Artist-Journalism</i> <i>The Guardian</i> United Kingdom

Source: created by the authors from the *curriculum vitae* of the experts

3. Results

3.1. Professional profiles associated to data journalism developed in Spain

Data journalism is a relatively recent discipline in the panorama of communication media, and although the digital transformation of journalism began at the end of the 20<sup>th</sup> Century, it was not until 2010 when the first data teams or departments were created in leading international media. In Spain, experts remember that in 2011, a data journalism group, Medialab



Matadero, emerged that offered free training to promote the discipline, “and the first initiative emerged from that point on” (E11).

More than a decade later, although journalists believe that the discipline has grown, and that the employment possibilities are real, it needs more development, looking at the more evolved English-speaking media with disappointment, where data “is not only associated to the present, but reports are also constructed with time and resources, unhurried stories outside the agenda” (E5).

“Little by little, but there is interest and more resources are increasingly provided”, indicates E4, in agreement with E8, who points out for the first time the three large areas of professional work: exploitation of data, verification, and visualization.

The experts describe these profiles:

- The data journalist is able to extract information from databases, interpret them, and analyze them to create a story “that goes beyond” (E3).
- The fact-checker journalist, the figure that experts consider the most cross-sectional, and who emerged during times of disinformation, is responsible for contrasting information with specific data programs to combat the proliferation of fake news.
- The visual journalist, who works more directly in the way in which the stories based on data are shown or told. This journalist creates maps, infographics, graphics, tables, and images that help in understanding the information.

E6 indicates that although not exactly in the area of data, web traffic analytics and metrics are indispensable in newsrooms. “Also, if the data (news) pieces work, and there is user interaction and following, it will help management invest in more resources”.

E12 explains the divisions created by the company she works for, which could provide us with clues about other professional areas that could be explored in the future:

- The audiovisual producer specialized on data, as all media tend to inform through images, the most consumed resource. He or she is responsible for the ideation and execution of a piece, but also the management of resources. The expert underlines the need to provide movement and spectacularization to the informative pieces, and knowing how to present data in a manner that is attractive, manageable, and aesthetic for the screen.
- The expert in new narratives, with the capacity to apply new formats to journalistic projects, through the incorporation of the latest trends in audience interactivity, gamification, and videogames. Likewise, they are granted with the ability to create reports based on mobile journalism, Web 2.0, and social networks.
- The Artificial Intelligence (AI) applied to journalism specialist, an expert who uses programs such as Deep Learning to perform automatic transcriptions of interviews and declarations to detect verifiable phrases from the transcriptions. He or she also uses tools such as Claim Hunter, which monitors Twitter to automatically detect verifiable messages. IA allows increasing the speed of checkers by 50%.

E8 supports the growth of positions linked with data journalism and the use of new narratives to enrich the experience of the user. This is how he or she determines how closely the *podcast* or video will be attached to the presentation of data. As well as the transmedia narrative and interactive visualizations, indicates E7.

In line with this, E9 reminds us that journalism now competes with Netflix, HBO, or Amazon Prime Video, and that “people take their phone and look for entertainment, so that data journalism, along with new narratives, will allow offering more attractive stories”.

Data journalism transcends the work in communication media. The experts point out that data analysts are becoming more valued in corporate journalism, where the communicators must know how to provide information about a company and create *storytelling* for the brand, and in the communication consultancies as well, where the mastery of exploitation and visualization of data tools is much appreciated (E9 and E10).

For the experts, the basic skills of a journalist must be joined to other more technical know-how such as access to publicly-available information, the functioning of databases, downloading and processing, optimized searches in Google, the use of visualization tools, and knowledge on the principles of graphic design, *scraping*, or *big data*. Likewise, although everyone believes that technology is fleeting, and that new tools emerge daily, they indicate that communication media outlets value, when hiring a journalist, a professional who has skills in the management of digital tools related to data journalism, and that imply actions such as: exploitation of databases (most of them point to Excel as the starting point, or OpenRefine), data analysis, and visualization tools (Tableau, After Effects, Flourish, or Infogram), and verification (Invid, Hotsuite, Crowd Anger).

With respect to the commitment of the media, all of them coincide that the health crisis due to COVID-19 was a turning point. The public needed to understand the information, and the data journalist specialization demonstrated its value. E10 exemplifies it with this comment: Readers understood what a basic health area was in Madrid, and were able to know the incidence (of the disease) in the municipality”.

E4 completes this topic, and details that the interest and development of the specialization was due “as a consequence of the explosion in data, through the different channels that offered information, from the ministry and up to the autonomous communities”. For this, “the media outlets were forced to look more seriously into the management of data, to provide daily information on the evolution of the impact of the pandemic”.

As for the re-organization of the newsrooms, all of them coincided in that the number of data journalists hired for this role is still low, and that when looking at these numbers, we cannot discuss an adjustment in work processes. However, we are starting to see the hybridization of the newsrooms, now composed by different profiles of journalists, as they share their workspace with programmers or designers.

Aside from the integration of technical teams, E11 provides an example from the English-speaking press, which is starting to hire “geographers or cartographers, and of course, *data scientists*”. “In the end, the journalist becomes a *product manager*, as it is he or she who knows all the approaches, the journalistic stories”, explains, “and the integrated teams are the ones who show the most resources”.

The reflection by E2 is very interesting, as it refers to the slowing down of the evolution of data journalism in Spain, given the difficulty in finding prepared professionals. “The current supply of this type of specialist is still scarce, specifically by the technical components needed, as there is still a dichotomy in the journalism world between letters and sciences, and in this new environment, communicators must utilize and manage numbers”. E10 adds another difficulty, the lack of investment in

technology: “Technology and specific programs form part of the management of resources when implementing or increasing the areas of data”.

### 3.2. Areas of data professionals in leading international media

Throughout the *focus groups*, the twelve journalists consulted pointed out that English-speaking media clearly led the development of data journalism in communication media worldwide. The examples of the United States and the United Kingdom were mentioned, and especially the examples of *The Washington Post*, *The New York Times* and *The Guardian*, respectively, so that it is interesting to count with the direct intervention of experts from these outlets.

In fact, E11 even affirms that in Spain, “the media outlets that are truly committing to the visualization and analysis of data, and that also promote the creation of interactive reports, are following the trail of these international media”.

In both markets, United States and United Kingdom, the professionals confirm that the newsrooms have their own data management teams, normally divided into two sub-areas: data, and graphics.

On the other hand, although international journalists point to specific profiles that are already developed, and which have a consolidated designation (data editors, data analyst, statistical journalist, expert in creation and use of polls, developer or designer journalists, among the most sought”, they coincide with the Spanish ones in the three great areas of activity: profiles associated to investigative journalism, fact-check professionals, and data journalists, who come from the areas of infographics, and who are visual artists.

In the case of the *Washington Post*, the area that is rapidly growing is graphics. The interviewee, Chiqui Esteban, has a Journalism degree from the University of Navarre, and currently directs the graphical department of this prestigious journal with a team of forty people under his management. He was previously associated with *National Geographic* and *The Boston Globe*, the latter being the media outlet in which he obtained the Pulitzer Prize and the Online Journalism Award for his coverage of the Boston bombing.

He confirms that despite having two differentiated teams, collaboration is frequent between *Data* and *Graphics*, and that “data are also collected from the latter to create their own stories”. The graphics part is increasingly growing because “the medium is very clear in that everything that can be done graphically will be done”. “The main stories add graphics and annotations; a lot of emphasis is placed on finding differentiated formats, creating pieces that are durable”, he ensures.

Esteban explained that since 2020, data journalism has gained more importance in his media outlet, and that the naming of Meghan Hoyer as data journalism director of the newspaper was very relevant. Her first act was to develop methodologies and workflows between the numerous journalists in the newsroom who were data specialists for developing investigative journalism, visual narrations, and reports centered on data.

The journalist details, as other experts who work in Spain anticipated, that in their medium, although data is present in “their day-to-day”, a commitment was made for “the forecasting of original topics”. “Our stories, the fast ones, take three to four days, and for others, we can work on it between one or two weeks. And, although we cover the news cycle, it doesn’t have to be necessarily the same day”.

He also explains how, when dealing with elaborate pieces, the organization of work is multidisciplinary, because experts from different fields are needed, underlining that hybrid teams are adequate for creating data-based departments. “For the large research projects, we count on different journalists; in the end, many of the articles are written by three or four authors”, he points out.

When a new team member must be hired, they evaluate certain skills. “It is indispensable, at least, to master spreadsheets, aside from the use of visualization tools”. Among the tools, he names Tableau, Flourish, or Datawrapper, “which do not require programming, although they have a learning curve”. “If, in addition, we add code management skills (programming), or *web scraping* to search the web, the profile is complete”.

In any case, he comments that in *The Washington Post*, the department is large, that “we don’t have to demand all of these skills from one person”. Nevertheless, he underlines a vital condition that asks for reflection: in this journal, there is a *sine qua non* that all of them be journalists. “They work in a media outlet, they will develop an article, and they have to ask the correct and adequate questions”, he reveals.

In the case of *The Guardian*, it must be underlined that in the *focus group*, it was remembered how Sept 11<sup>th</sup> was crucial for the development of a data department in this media outlet, and that Simon Rogers, data editor, still states, 10 years after, that the attack showed them that stories could be told not only through words, but also graphics, in a visual manner.

For Pablo Gutiérrez, *Visual Artist-Journalism* at *The Guardian*, education and specialization in data “without a doubt” improves the employability of the journalist at present. Although educated at the Complutense University of Madrid in the field of arts, he has always worked, however, in communication; first, in the creation of interactive graphics, and infographics at *El Mundo*, and since 2013, in international media such as the BBC before incorporating to his current media outlet in 2015.

In the British newspaper, the data and graphics departments are also independent: “They are separated, but we tend to establish some collaborations during the year depending on the story, if it can be explained visually”.

In his case, he is specialized in visually narrating these stories when a dynamic or interactive component is needed. “My work is generally related with the use of cartography, *storytelling*, and data visualization”, he says, when explaining that in the *Visuals* Department, the professional profiles “are very diverse”. “They come from degrees in Journalism, Politics, Programming, Design, and Arts”, he specifies, as opposed to the American medium analyzed. Also, very specific specialists are hired within the visuals part, such as experts in “Git, Java Script (JS), AWS, Lambdas, D3.js, Geographical Information System (GIS), design tools, and 3D or back-end and front-end programming”.

In fact, he highlights that these latter profiles are the most difficult to find, and that it is “strange to find a hybrid professional between development, journalism, and design”.

#### 4. Discussion and conclusions

Data journalism has become one of the most important components of information reporting in the last few years, but its evolution has not advanced at the same pace in Spain as in the English-speaking world. The professionals consulted detected important progress in specific media, corporations, and entities in Spain, but identified two circumstances that stop its

development: the lack of qualified professionals-further extending what was already discussed by Cherubini et al. (2021) and Newman (2022), and the insufficient investment by news agencies.

In Spain, there are specific academic degrees to teach skills to future data journalists, especially at the graduate level, but most of the degrees, especially those that are official, have only been implemented in the last few years. Specific Spanish media outlets have added important professionals to their staff, who have become leaders in Spain. However, most of the newsrooms have not received the investment necessary, neither the personnel nor the infrastructures, to be able to create a team specialized on data and permanently focused on this type of work.

Those consulted coincided in that the health crisis due to COVID-19 was a turning point, as the public needed to understand the information, and data journalism demonstrated its value. These results are in agreement with the results from recent reports, such as that published by the Reuters Institute, or the ONTS. Likewise, they demand the application of data journalism not only in the present, considering that the use of techniques from this discipline to other, more relaxed genres, such as reports, would enable journalists to seduce a broader audience, and stress the necessary professionalization of the work, as in leading English-speaking media.

The surge of data, for which the context of the pandemic was a major boost, accelerated processes that had a great impact on the functioning of newsrooms, where the teams have an increasingly marked multidisciplinary and technical character. In this sense, our study aligns with those by Salaverría-Aliaga and Martínez-Costa (2021), Gómez-Calderón et al., (2017), Heredero-Díaz and Reyes-Sánchez (2016), Manfredi-Sánchez et al., (2019) and González-Alba (2022).

The new professional profiles are associated with three areas that are particularly flourishing: visualization, fact-checking, and investigative journalism. The rise of the first was pointed out by López-García et al. (2016), among others, while Cerezo-Gilarranz (2020) highlighted the second.

But also, journalists participate in the objective of creating transmedia stories, with a notable audiovisual component, creating attractive stories that capture the attention of the audience. Thus, the profile of the audiovisual producer specialized in data becomes settled, along with the expert in new narratives, who adds trends in interactivity, gamification, and videogames. The areas of development are completed by the specialist on AI applied to journalism, who increases the speed of the work of the fact-checkers.

The professionals believe that the application of data will increase in *podcasts* and videos, and that the transmedia narrative and interactive visualizations of information can contribute towards journalism, which is also an agent of entertainment, being able to compete with the great global platforms, which are coming closer to non-fictional contents.

In the Spanish media, employability increases for journalists with skills on data mining, data analysis, fact-checking of content, and creation of infographics. And with the addition of knowledge on programming or statistics, we find ourselves with a highly qualified profile.

In the English-speaking media consulted, the teams who worked with data were numerous, and the journalist habitually interacted with experts in informatics, Artificial Intelligence, design, or mathematics. In general, they were organized into two interconnected departments: data and graphics, with a higher growth observed in the latter. The teams worked on quality

pieces, beyond the immediacy dictated by the news, with many days spent on significant pieces, which follow the maxims of quality journalism.

In the international media referenced, it was observed that in integrated teams, the journalist starts to become a *product manager*, coordinating the functions of the different professionals, as he or she knows the focus that the story must follow, and the characteristics of the audience.

The profile profiles associated with programming are the most demanded, as these are professionals who are currently in high demand in many spheres. Thus, the talents associated with the exploitation and interpretation of data broaden their employment possibilities, and the offers are not only centered on media, but the ability to provide data for making decisions is a guarantee for success in any sphere.

Therefore, data journalism is consolidated within the media ecosystem, as a very interesting aspect, which offers new options for professional development, and with an essential role in a society that is facing severe challenges such as disinformation, where the massive verification of data emerges as an indispensable task for media, agencies, and independent entities.

## 5. Acknowledgements

The present article was translated by Mario Fon, whom I thank for his work.

Sincere thanks to all the data journalism professionals in Spain and other countries, who have generously shared their experience and knowledge to the authors of the present study.

## 6. Specific contributions by each author

	Name and surname
Conception and design of the work	Marta Saavedra, Mercedes Herrero and Eduardo Castillo
Methodology	Marta Saavedra
Data collection and analysis	Marta Saavedra, Mercedes Herrero and Eduardo Castillo
Discussion and Conclusions	Mercedes Herrero
Drafting, formatting, review and approval of the different versions	Marta Saavedra, Mercedes Herrero and Eduardo Castillo

## 7. Bibliographic references

- Amoedo-Casais, A., Vara-Miguel, A., Negredo-Bruna, S., Moreno-Moreno, E. y Kaufmann-Argueta, J. (2022). *Digital News Report España 2022*. Pamplona: Reuters Institute for the Study of Journalism, Universidad de Navarra. <https://doi.org/10.15581/019.2022>
- Bauman, Z. (2005). *Vida líquida*. Madrid: Paidós Ibérica.
- Casal-Cortizas, F.M. (2010). *El periodismo de investigación asistido por ordenador y de precisión aplicado a la prensa local y regional: el contraste entre Estados Unidos y un modelo de futuro para Galicia*. [Tesis doctoral]. Universidad Complutense de Madrid. <https://cutt.ly/U052dwo>
- Cerezo-Gilarranz, P. (27 mayo 2021). *Modelos de negocio. Desafíos de la transformación digital*. XXVII Congreso de la Sociedad Española de Periodística (SEP), Universidad de Sevilla-SEP, Sevilla.
- Chaparro Domínguez, M.A. (2013). *La evolución del periodismo de precisión: el blog de The Guardian sobre periodismo de datos*. Actas del I Congreso Internacional Comunicación y Sociedad, UNIR, Logroño. <https://cutt.ly/CXdJvoe>
- Cherubini, F., Newman, N. y Nielsen, R.K. (2021). *Changing Newsrooms 2021. Hybrid Working and Improving Diversity Remain Twin Challenges for Publishers*. Oxford: Reuters Institute for the Study of Journalism-University of Oxford. <https://cutt.ly/JXhHoft>
- Córdoba-Cabús, A., García-Borrego, M. y López-Martín, A. (2020). El periodismo de datos durante la crisis sanitaria del Covid-19 en la prensa española. *Revista Iberica de Sistemas e Tecnologías de Informação*, E35, 325-337. <https://cutt.ly/wnZtjhq>
- Dader-García, J.L. (1997). *Periodismo de precisión. Vía socioinformática de descubrir noticias*. Madrid. Síntesis.
- Dader-García, J.L., Gómez-Fernández, F.P. (1993). Periodismo De precisión: Una Nueva Metodología Para Transformar El Periodismo. *Anàlisi: Quaderns De Comunicació I Cultura*, 15, 99-116. <https://cutt.ly/H052mKW>
- DeFleur, M. (1997). *Computer-Assisted Investigative Reporting. Development and Methodology*. Mahwah: Erlbaum (LEA).
- Flores-Vivar, J. y Salinas-Aguilar, C. (2013). El periodismo de datos como especialización de las organizaciones de noticias en Internet. *Correspondencias & Análisis*, 3, 15-34. <https://doi.org/10.24265/cian.2013.n3.01>
- Edmunds, H. (1999). *The focus group research handbook*. Chicago: NTC/Contemporary Publishing Group.
- Ferreras-Rodríguez, E.M. (2016). El periodismo de Datos en España. *Estudios sobre el Mensaje Periodístico*, 22(1), 255-272. [http://dx.doi.org/10.5209/rev\\_ESMP.2016.v22.n1.52594](http://dx.doi.org/10.5209/rev_ESMP.2016.v22.n1.52594)
- Ferreras-Rodríguez, E.M. (2013). Aproximación teórica al perfil profesional del “Periodista de datos”. *Revista ICONO 14. Revista Científica de Comunicación y Tecnologías Emergentes*, 11(2), 115-140. <https://doi.org/10.7195/ri14.v11i2.573>
- Fink, K. y Anderson, C.W. (2014): Data Journalism in the United States. *Journalism Studies*, 16(4), 467-481. <http://dx.doi.org/10.1080/1461670X.2014.939852>
- García-Avilés, J.A. (2021). Review article: Journalism innovation research, a diverse and flourishing field (2000-2020). *Profesional de la información*, 30(1). <https://doi.org/10.3145/epi.2021.ene.10>

García-Avilés, J.A., Carvajal-Prieto, M., Arias-Robles, F. y De Lara-González, A. (2019). Journalists' views on innovating in the newsroom: Proposing a model of the diffusion of innovations in media outlets. *The Journal of Media Innovations*, 5(1), 1-16. <http://dx.doi.org/10.5617/jomi.v5i1.3968>

Gavin, H. (2008). *Understanding research methods and statistics in psychology*. Londres: Sage Publications.

Gehrke, M., y Mielniczuk, L.P. (2017). Philip Meyer, the outsider who created Precision Journalism. *Intexto* 39, 4-13. <http://dx.doi.org/10.19132/1807-8583201739.4-13>

Gómez-Calderón, B., Roses-Campos, S. y García-Borrego, M. (2017). Los nuevos perfiles profesionales del periodista desde la perspectiva académica española. *Revista Mediterránea de Comunicación*, 8(1), 191-200. <https://www.doi.org/10.14198/MEDCOM2017.8.1.14>

González-Alba, J.A. La adaptación de los profesionales a la transformación digital de los medios: los nuevos perfiles que se incorporan a las redacciones. *Cuadernos de periodistas*, 23 febrero 2022. <https://cutt.ly/lXjPjUX>

Herederó-Díaz, O. y Reyes-Sánchez, F. (2016). Los efectos profesionales de la crisis de la televisión pública española: el creador audiovisual low cost y las nuevas tecnologías. *AdComunica. Revista Científica de Estrategias, Tendencias e Innovación en Comunicación*, 11, 95-118. <http://dx.doi.org/10.6035/2174-0992.2016.11.7>

Herrero-De-la-Fuente, M. y Ríos-Calvo, C. (2022). Construcción de un escenario para la posverdad. Redes sociales y desinformación. En A. Pérez-Escoda y J. Rubio-Romero (eds.). *Redes Sociales, ¿el quinto poder? Una aproximación por ámbitos al fenómeno que ha transformado la comunicación pública y privada* (pp. 81-97). Valencia: Tirant lo Blanch.

Hewett, J. (2015). Learning to teach data journalism: Innovation, influence and constraints. *Journalism*, 17(1), 119-137. <https://doi.org/10.1177/1464884915612681>

Kotler, P. y Armstrong, G. (2016). *Marketing*. Reading (MA): Addison Wesley.

La Rosa-Barrolleta, L. (2019). *Periodismo de datos: el big data como elemento diferenciador sociocultural e insuficiencia de la Ley de Transparencia en España*. [Tesis doctoral]. Universidad Carlos III (Madrid). <https://cutt.ly/g053dun>

López-López, P.J. (2016). *Antecedentes y fundamentos del Periodismo de datos*. I Congreso Internacional sobre Periodismo de Datos. Nuevas Narrativas para el Periodismo Especializado, Universidad de Málaga, Málaga. <https://cutt.ly/D059G5M>

López-García, X., Toural-Bran, C. y Rodríguez-Vázquez, A.I. (2016a). Software, estadística y gestión de bases de datos en el perfil del periodista de datos. *El profesional de la información*, 25(2), 286-294. <http://dx.doi.org/10.3145/epi.2016.mar.16>

López-García, X., Rodríguez-Vázquez, A.I. y Álvarez-Gromaz, L. (2016b). Análisis de las experiencias de 'The Washington Post' y 'The Guardian'. *Revista Telos*, 103, 1-14. <https://cutt.ly/2XdJLqH>

López-García, X. (2001). Nuevos perfiles de los periodistas en la sociedad de la información. *Ámbitos. Revista Andaluza de Comunicación*, 78, 7-18. <https://cutt.ly/EXjF9hj>

Manfredi-Sánchez, J.L., Ufarte-Ruiz, M.J. y Herranz-De la Casa, J.M. (2019). Innovación periodística y sociedad digital: Una adaptación de los estudios de Periodismo. *Revista Latina de Comunicación Social*, 74, 1633-1654. <https://doi.org/10.4185/RLCS-2019-1402>



- Meyer, P. (2002). *Precision Journalism. A Reporter's Introduction to Social Science Methods*. Lanham: Rowman and Littlefield.
- Nee, Rebecca C. (2013). Creative destruction: An exploratory study of how digitally native news nonprofits are innovating online journalism practices. *International journal on media management*, 15(1), 3-22. <https://doi.org/10.1080/14241277.2012.732153>
- Newman, N. (2022). *Journalism, Media, and Technology Trends and Predictions in 2022*. Reuters Institute for the Study of Journalism-University of Oxford. <https://cutt.ly/3XhHWEe>
- Picard, R.G. (2017). *Value Creation and the Future of News Organizations. Why and how journalism must change to remain relevant in the twenty first century*. Oporto: Media XXI.
- Rodríguez-Martínez, R. (2021). *Desinformación y plataformas de fact-checking: estado de la cuestión*. Barcelona: Universitat Pompeu Fabra. <https://cutt.ly/wXje5tv>
- Rubio-Jordán, A.V. (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. Revista Interdisciplinar De Estudios De Comunicación Y Ciencias Sociales*, 32, 289-303. <https://doi.org/10.31921/doxacom.n32a14>
- Saavedra-Llamas, M. y Toledano Cuervas-Mons, F. (2022). La comunicación como estudio transversal: evolución de las facultades especializadas en España. *Revista de Comunicación de la SEECI*, 55, 211-231. <http://doi.org/10.15198/seeci.2022.55.e799>
- Saavedra-Llamas, M., Herrero-De-la-Fuente, M. y Castillo-Lozano, E. (2020). La formación en periodismo de datos en España: radiografía de la oferta académica universitaria. *Anàlisi: Quaderns de Comunicació i Cultura*, 62, 93-109. <https://cutt.ly/PVWolq1>
- Saavedra-Llamas, M., Herrero-De-la-Fuente, M., Rodríguez-Fernández, L. y Jiménez-Narros, C. (2019). Información de salud: fuentes periodísticas y desafíos profesionales. *El profesional de la información*, 28(2). <https://doi.org/10.3145/epi.2019.mar.08>
- Salaverría-Aliaga, R. y Martínez-Costa, M.P. (eds.) (2021). *Medios nativos digitales en España: caracterización y tendencias*. Salamanca: Comunicación Social.
- Salvat-Martinrey, G. y Aranda-Martínez, S. (2020). Análisis de contenido de la prensa diaria española sobre el uso de la infografía para informar de la crisis sanitaria de la COVID-19. *Revista de Comunicación y Salud*, 10(2), 185-210. [https://doi.org/10.35669/rcys.2020.10\(2\).185-210](https://doi.org/10.35669/rcys.2020.10(2).185-210)
- Sánchez-Marín, G. y Micó-Sanz, J.L. (2014). Perfiles profesionales en las agencias de noticias. Estudio comparativo de ACN, EFE y Europa Press. *El profesional de la información*, 23(5), 501-509. <http://dx.doi.org/10.3145/epi.2014.sep.07>
- Scolari, C.A. (13 agosto 2021). Adiós sociedad líquida. Bienvenida sociedad gaseosa. *Hipermediaciones*. <https://cutt.ly/fXhu6IG>
- Valles-Martínez, M.S. (2002). *Cuadernos metodológicos. Entrevistas cualitativas*. Madrid: Centro de Investigaciones Sociológicas.

Vázquez-Medel, M.A. y Pulido-Rodríguez, C. (2021). Más allá de la desinformación y las *fake news*. En L.M. Pedrero-Esteban y A. Pérez-Escoda (eds.). *Cartografía de la comunicación postdigital. Medios y audiencias en la sociedad de la Covid-19* (pp. 201-222). Pamplona: Thomson Reuters-Aranzadi Civitas.

Velasco-Jones, L., Muñoz-López, L., Antón-Martínez, P. y Cáceres-Piñuel, M. (eds.) (2021). *Uso y actitudes de consumo de contenidos digitales en España, 2021*. Madrid: Observatorio Nacional de Tecnología y Sociedad-Ministerio de Asuntos Económicos y Transformación Digital. <https://doi.org/10.30923/094-21-023-9>

Vivar-Zurita, H. (2011). Nuevos perfiles profesionales para una comunicación digital. *Revista Telos*, 87, 58-62. <https://cutt.ly/8XjGWTr>

Vizoso-García, Á. (2021). Periodistas de datos y visualización de la información. En R. Salaverría-Aliaga y M.P. Martínez-Costa (eds.). *Medios nativos digitales en España: caracterización y tendencias* (pp. 247-256). Salamanca: Comunicación Social.

Zapater-Delgado, P. (2021). *La transformación digital en The Washington Post: estudio cualitativo de la relación entre periodismo y tecnología en la era de Jeff Bezos*. [Tesis doctoral]. Universidad de Zaragoza. <https://cutt.ly/i0ZljWq>