Data Journalism in the Spanish Caribbean Digital Media*

Ramaris Albert Trinidad
Universidad de Puerto Rico. Escuela de Comunicación
ramaris.albert@upr.edu

Submission date: June 2019
Accepted date: May 2020
Published in: October 2020

Recommended citation: ALBERT TRINIDAD, R. (2020). "Data Journalism in the Spanish Caribbean Digital Media". *Anàlisi: Quaderns de Comunicació i Cultura*, núm. extraordinari 2020, 47-66. DOI: <https://doi.org/10.5565/rev/analisi.3236>

Abstract

21st-century journalism is characterized by new trends that facilitate the production of digital and novel contents. Among them, most notable is data-driven journalism, which collects structured information and analyzes it in-depth to present it in more versatile formats through interactive visualizations and other multimedia tools. This specialty serves to tell stories based on vast amounts of information that, without the assistance of a computer and the application of statistical methods, would be challenging to report. This work is a pilot study about the degree of penetration of data journalism through a comparative study of the journalistic media of the Spanish-speaking Caribbean. For this purpose, an exploratory pilot questionnaire investigated the processes and techniques used by journalists in the preparation of publications in digital media of Puerto Rico, Cuba, and the Dominican Republic, in order to determine whether they apply data journalism from the defining elements that Paul Bradshaw establishes in his inverted pyramid of data journalism. The findings pointed out that journalists from these countries vary in the level of frequency of use of the five characteristic functions, of which the analysis of databases is the most used by all. The respondents agree that the least used data journalism techniques are the collection of data through contests or open collaboration projects (crowdsourcing) and the development of tools and informative web applications based on database contents. The digital journalists who have trained in this discipline did so, mainly, through workshops and practical seminars they took on their own initiative.

Keywords: data journalism; digital media; Cuba; Dominican Republic; Puerto Rico

* Research carried out through the pilot project of Resident Academics of the Deanship of Graduate Studies and Research (DEGI) with the support of the Center for Communication Research (CICom) of the School of Communication of the University of Puerto Rico.

Resum. *Periodisme de dades en els mitjans digitals del Carib hispà*

El periodisme del segle XXI es caracteritza per noves tendències que possibiliten la producció de continguts digitals i nous. Entre aquestes, destaca el periodisme de bases de dades, especialitat que ofereix informació, context i anàlisi amb major profunditat i en formats més versàtils a través de les visualitzacions interactives i altres eines multimèdia. Aquesta especialitat permet contar històries a base de grans quantitats d'informació que, sense l'assistència d'una computadora i l'aplicació de mètodes estadístics, serien molt difícils de reportar. Aquest treball proposa un estudi sobre el grau de penetració en què es troba l'exercici del periodisme de dades mitjançant un estudi comparatiu dels mitjans periodístics del Carib hispanoparlant. Per fer-ho, un qüestionari pilot exploratori va investigar els processos i les tècniques utilitzades pels periodistes en l'elaboració de publicacions en mitjans digitals de Puerto Rico, Cuba i la República Dominicana, amb la finalitat de determinar si apliquen el periodisme de dades a partir dels elements definitoris que estableix Paul Bradshaw en la seva piràmide invertida del periodisme de dades. Els resultats apunten al fet que els periodistes dels tres països varien en el nivell de freqüència d'ús de les cinc funcions característiques, de les quals l'anàlisi de bases de dades resulta ser la més utilitzada per tots. Els enquestats coincideixen que les tècniques de periodisme de dades menys utilitzades són la recopilació de dades mitjançant concursos o projectes de col·laboració oberta (*crowdsourcing*) i el desenvolupament d'eines i aplicacions web informatives a partir de continguts de bases de dades.

Paraules clau: periodisme de dades; mitjans digitals; Cuba; República Dominicana; Puerto Rico

Resumen. *Periodismo de datos en los medios digitales del Caribe hispano*

El periodismo del siglo XXI se caracteriza por nuevas tendencias que posibilitan la producción de contenidos digitales y novedosos. Entre ellas, destaca el periodismo de bases de datos, especialidad que ofrece información, contexto y análisis con mayor profundidad y en formatos más versátiles a través de las visualizaciones interactivas y otras herramientas multimedia. Esta especialidad permite contar historias a base de grandes cantidades de información que, sin la asistencia de una computadora y la aplicación de métodos estadísticos, serían muy difíciles de reportar. Este trabajo propone un estudio sobre el grado de penetración en el que se encuentra el ejercicio del periodismo de datos mediante un estudio comparativo de los medios periodísticos del Caribe hispanoparlante. Para ello, un cuestionario piloto exploratorio investigó los procesos y las técnicas utilizadas por los periodistas en la elaboración de publicaciones en medios digitales de Puerto Rico, Cuba y la República Dominicana, con el fin de determinar si aplican el periodismo de datos a partir de los elementos definitorios que establece Paul Bradshaw en su pirámide invertida del periodismo de datos. Los resultados apuntan a que los periodistas de los tres países varían en el nivel de frecuencia de uso de las cinco funciones características, de las que el análisis de bases de datos resulta ser la más utilizada por todos. Los encuestados coinciden en que las técnicas de periodismo de datos menos utilizadas son la recopilación de datos mediante concursos o proyectos de colaboración abierta (*crowdsourcing*) y el desarrollo de herramientas y aplicaciones web informativas a partir de contenidos de bases de datos.

Palabras clave: periodismo de datos; medios digitales; Cuba; República Dominicana; Puerto Rico

1. Introduction

For several years, the media have been struggling against a doubly adversarial scenario: an international economic crisis which led to a decline in advertising revenues, and a change in the information consumption habits of citizens, forcing them to transform their work model towards digital platforms as the primary support (Salaverría, 2015).

Both circumstances have shown the need to transform journalistic companies in order to find formulas for success in the creation of innovative content, especially in digital platforms. This search for new ways of informing points internationally to multimedia innovation, digital visualizations, and data journalism as some alternatives (Salaverría, 2015).

In this context, this article reveals data journalism in the digital media of the Spanish-speaking Caribbean as a booming specialty that provides added value to journalistic content.

The general objective of this research was to evaluate the degree of presence and penetration of data journalism in the performance of digital media through a comparative study of the Spanish-speaking Caribbean countries. For this purpose, it met the following specific objectives:

1. Identify which data journalism techniques are used in the digital media by journalists in Puerto Rico, Cuba, and the Dominican Republic.
2. Compare how frequently journalists in the three countries use data journalism techniques.
3. Determine if digital journalists from these countries have received professional training aimed at creating data journalism content and if so, which person or entity developed this initiative (for instance, personal or enterprise-driven).

The following research questions will be answered through the achievement of the aforementioned objectives:

1. What are the main data journalism techniques incorporated by journalists in the digital media in Puerto Rico, Cuba, and the Dominican Republic?
2. What are the data journalism techniques that are least used by journalists in digital media in Puerto Rico, Cuba, and the Dominican Republic?
3. Do journalists receive any professional training in their media aimed at creating data journalism content?
4. Has the professional training that journalists have followed emerged as their own initiative or been driven by the interests of the enterprise for which they work?

In accordance with the research carried out, this article has an exploratory scope since it conducts, for the first time, a diagnosis of the techniques and tools currently used in the development of digital journalistic products of the media in these three countries, from the perspective of data journalism. This publication broadens the international knowledge of a specialty that has been little explored to date in the Hispanic Caribbean and the presence it has in these locations.

From a practical perspective, the results of this study allow the media to determine creative skills and useful tools in the development and publication of multimedia content. This does not only help media to decipher effective strategies to face the challenges of offering innovative content, but it also contributes to the establishment of sustainable functioning models. Furthermore, it also results in a different, more vibrant, and relevant journalistic service for a society that is continuously increasing its information consumption through digital platforms.

2. Literature review

Data-driven journalism is a specialty that allows journalists to report stories based on the collection and analysis of large amounts of information that, without the assistance of a computer and statistical methods, would be very difficult to report (Bradshaw and Rohumaa, 2011). This allows us to offer information, context, and analysis in greater depth and in more versatile formats through visualizations and other multimedia tools. This discipline gives the journalist the tools to handle structured information (Holovaty, 2006) and tell a complex story through interactive resources, helps explain how it relates to people or serves as a news collection process in itself (Bradshaw, 2013; Bradshaw and Maseda, 2015).

In 2006, journalist and web designer Adrian Holovaty said that "newspapers need to stop the story-centric worldview" (Holovaty, 2006: para.5) and "build an infrastructure that makes them reliable data centers, capable of analyzing even very large and complex data sets internally and build stories about their knowledge" (Holovaty, 2006: para.7). This approach was subsequently recognized as the manifesto of data journalism (Kayser-Bril, Valeeva and Radchenko, 2016) and is one of the first reflections on this new discipline, which is preceded by the Computer Assisted Reporting (Egawhary and O'Murchu, 2012) developed by journalist Philip Meyer through an analysis of the profile of protesters in the Detroit riots in 1967. Meyer had already dabbled in precision journalism (Crucianelli, 2013) a few years before investigating the high prices of school insurance in Florida (Meyer, 2012) and discovering, through data-crossing, that the insurers financed the campaigns of the members of the councils of officials who signed the policies (Chaparro, 2014). Other notable antecedents in the emergence of data journalism are the visualization works of French engineer Charles Joseph Minard in his "Figurative Map of the successive losses in men of the French Army in the

Russian campaign 1812–1813" and the English epidemiologist John Snow, who discovered that cholera was transmitted through water and created a visualization of those killed by the disease in the London district of Soho and the location of the wells from where they took the contaminated water (Kayser-Bril, Valeeva and Radchenko, 2016).

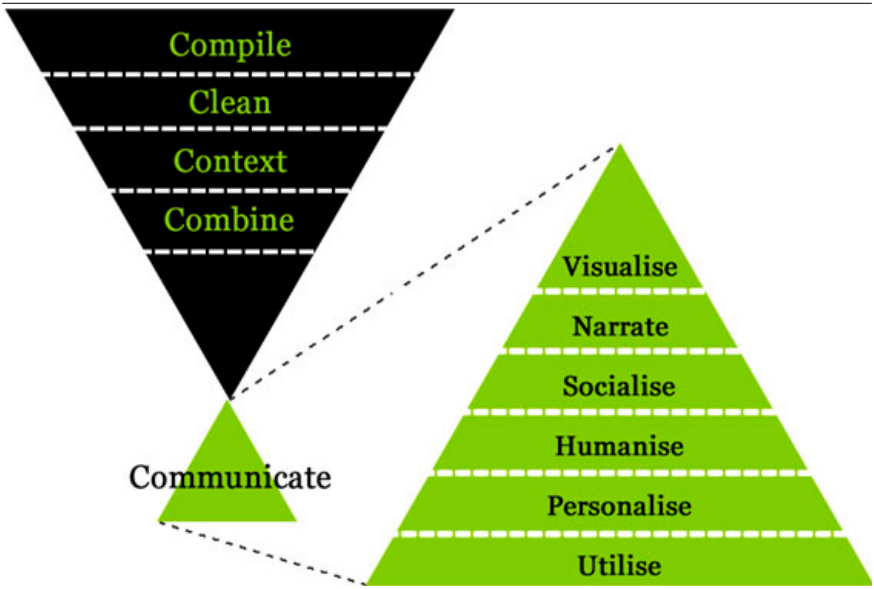
3. Theoretical framework

Following the Holovaty manifesto and the launch of the Guardian Datablog in 2009 (Chaparro, 2013), in 2011 Paul Bradshaw developed the term "inverted pyramid of data journalism," which is a structure that defines the phases needed to build a data journalism story. Bradshaw emphasizes that its application in this specialization is a "contrast to the classical inverted pyramid taught in any faculty of journalism." (Bradshaw, 2011a: para.1) For the author, instead of presenting the structure of a conventional text, "it begins with a large amount of information which becomes increasingly focused as you drill down into it until you reach the point of communicating the results" (Bradshaw, 2011a: para.2). This inverted pyramid of data journalism is composed of five phases: compile, clean, contextualize, combine, and communicate which, in turn, are broken down into visualise, narrate, socialise, humanise, personalise, and utilise. Each of these processes encompasses various techniques by which diverse contents of data journalism can be performed. This model serves to determine if a journalistic work was carried out following the phases needed in the exercise of data journalism (Figure 1).

The combination of reporting techniques with the integration of visualizations and the use of technology for the collection, exploitation, and analysis of data are part of this professional exercise, which must be "open, accessible, and enlightening" by publishing the data that inform the report in reusable formats (Rogers, 2016: para.5). Rogers agrees that "the compilation of data is what defines it as an act of data journalism" (Bradshaw, 2011a: para.5). Also, he highlights this action as a crucial element, since together with "the editorial idea behind this process, the idea that you are going to tell a story here," (Rogers, 2014: para.22) is what makes data journalism "different to just publishing the data" (Rogers, 2014: para.22).

The application of data journalism multiplies the possibilities of developing stories that appeal and connect with the topics that interest citizens. Other models of data journalism include techniques such as a shared database for journalists from different media, global collaboration between journalists, joint forces of journalists, activists, and technologists, crowdsourcing, open source intelligence, and databases for users to research (Kayser-Bril, Valeeva and Radchenko, 2016). This versatility, rigor, and the ability it affords readers to interact with its content have made it one of the most effective ways to add value to informative content through multimedia formats (Lorenz, 2016). Its analytical and investigative nature has generated a growing need in newsrooms for journalists capable of producing multime-

Figure 1. The Inverted Pyramid of Data Journalism (complete)



Source: Paul Bradshaw, <https://onlinejournalismblog.com/>.

dia stories through data analysis (Treadwell, Ross, Lee and Lowenstein, 2016). Since 2011, at least six Pulitzer Prizes have been awarded to reports based on data journalism techniques, including the Panama Papers (Hudson, 2017)[1].

In the United States, data journalism is sometimes employed as part of investigative journalism (Kaplan, 2013). Its analytical nature and depth are closely linked to the publication of reports that monitor abuse of power, administrative irregularities or social complaints (Rogers, 2016). This specialization has not only evolved from professional practice in the newsrooms (Gray, Bounegru and Chambers, 2012; Verborgh and De Wilde, 2013; Casares and Chivite, 2015 among others). Additionally, it appears as a subject in university academic programs specialized in Journalism[2] (Chaparro, 2014). Recently, in countries such as Spain (Salaverría, Martínez-Costa and Breiner, 2018), it has been incorporated in the editorial offices of some media (López and Ufarte, 2016), providing new information, angles, and ways of

1. Some of the award-winning reports are *Floridians at Risk* (Herald Tribune, 2010), *Methadone and The Politics of Pain* (The Seattle Times, 2011), *Speeding cops get punished* (Sun Sentinel, 2012), *Medicare Unmasked* (Wall Street Journal, 2014), *People Shot Dead by Police This Year* (The Washington Post, 2015) and *Panama Papers* (ICIJ, 2016).
2. Some of the American universities consulted that have courses in data journalism are American University, Indiana University Bloomington, University of Maryland, University of Minnesota, and University of Missouri, among others.

interacting with citizens through reports and multimedia specials (Salaverría, 2015).

In Latin America, digital journalism began at the end of the 1990s, with a varied growth rate, conditioned by the social, political, and economic environments of each country (Salaverría, Rivera-Rogel and González-Córdova, 2019). With the help of collaborators from 22 countries, Salaverría (2016) traces an in-depth analysis of the origin, evolution, and trends of digital media in Ibero-America, which gives an account of the history, technological context, professional profile, training, and legal framework of online journalism in three Caribbean countries that are being studied in this research: Puerto Rico, Dominican Republic, and Cuba.

In the field of social sciences and communication, different studies on these countries draw on the book *La isla que se repite: El Caribe y la perspectiva posmoderna*, by the writer and economist Antonio Benítez Rojo (1989). In his work, Benítez Rojo establishes the interconnections between Puerto Rico, Dominican Republic, and Cuba, tracing the construction of an inexistent island based on their cultural, political, and economic characteristics (Colón, 2017). Based on this work and on Surlin and Soderlund's (1990) descriptions of the mass media in the Caribbean in the context of the Cold War, Colón locates the analysis of the mass media in these three countries within a continuous process of transition and transformation towards democratization. Colón links this process to the current international forces of globalized neoliberal economic and financial structures and the cultural transformations that these entail (2017).

Puerto Rico gained access to the internet in 1991, in a joint effort between the government and the University of Puerto Rico. However, news agencies had been working with computers and teletypes since the 1980s. El Nuevo Día, the country's leading newspaper, was the first to launch its portal in 1996, and Noticel, the first digital native newspaper, emerged in 2011. González (2016) notes that, since founded in 2007, the *Centro de Periodismo Investigativo* is the only independent media center responsible for conducting investigative journalism, employing data analysis techniques that are similar to data journalism. In August 2019 laws 122 and 141 came into effect, respectively known as the Open Data Law of the Government of Puerto Rico and the Law of Transparency and Expedited Procedure for Access to Public Information (Lexjuris, 2019). Their content has been objected to by journalistic associations, media, civil society organizations, and members of the legal class (Alvarado, 2019).

Network connection arrived in Dominican Republic in 1995. The Dominican Electronic Journal —its first digital media— was founded in March of the next year followed by the electronic version of the oldest printed newspaper in the country, Listín Diario (Ortiz, 2016). Like the rest of Latin American countries, the first stages of digital journalism in the Dominican Republic were characterized by the news dump of the editions printed on the electronic portals. However, Ortiz points out that the country is inte-

grating social media into the media scenario, including the mass use of resources such as Twitter, Facebook, and YouTube in the coverage and dissemination of journalistic content. "It is a moment that coincides with the boiling of the phenomenon of blogs or binnacles in the country, from the beginning of 2008, a universe from which Dominicans had remained relatively aloof until then" (Ortiz, 2016: 370). Cuba was officially connected to the internet in 1996. Granma Internacional, launched that same year, was the first Cuban newspaper in the network. According to Darias and Suárez (2016), the development of digital journalism in Cuba has been closely linked to government attempts to politically take advantage of new information technology. This is one of the reasons to which the authors attribute the fact that some cybermedia took until 2009 to begin to include interactive tools such as readers' comments and surveys, characteristic of digital journalism that Web 2.0 brought with it (Darias and Suárez, 2016: 122). As recently as December 6, 2018, Cuba activated the internet service through mobile data through the Empresa de Telecomunicaciones de Cuba (Antón, 2018). However, this country does not yet have a law or an official open data portal, such as the Dominican Republic —whose portal <https://datos.gob.do/> started in 2015— or Puerto Rico, which owns the government website <https://data.pr.gov/>, as well as the data inventory of the Institute of Statistics of Puerto Rico on the web <https://estadisticas.pr/>.

Researchers believe that Cuba is carrying out an intensive introduction of information and communication technologies in society, through the development and deployment of information systems to increase the efficiency of various economic and social sectors (Hidalgo-Delgado, Mariño-Molerio, Amoroso-Fernández and Leiva-Mederos, 2018). However, the opening of linked open data requires the use of appropriate mechanisms related to an updated technological infrastructure, an adequate legal framework, and adequate synergy between the producers and consumers of the open data linked in the country, which has not yet materialized.

After years of research, digital media in Latin America is described as a fundamental part of the newspaper market:

> Many of these media have adopted a general information approach, aimed at a heterogeneous audience. They also respond to quality parameters, including accessibility, visibility, and popularity. They are also subject to an important parameter: the geographical location, which is always limited by the levels of access to the network. (Salaverría, Rivera-Rogel and González-Córdova, 2019: 277)

The degree of presence and use of data journalism as a technique in the Spanish-speaking Caribbean is still unknown, since, to date, no research on data journalism has been published in Puerto Rico, Cuba, or the Dominican Republic. However, these advances towards greater accessibility to consume multimedia contents give greater relevance to studies on the processes of cre-

ating journalistic content on digital platforms, with data journalism as one of the trends with the highest growth in the field.

4. Methodology

This research was conducted with a quantitative approach. The measurement instrument was an exploratory pilot questionnaire including closed questions about the processes and techniques applied by digital media journalists in the selected countries. Drawing on the analysis of the five stages detailed by Bradshaw in his inverted pyramid of data journalism, a total of 15 techniques were included in the questionnaire, using a Likert frequency scale. The participants indicated how regularly these techniques apply to their work by answering always, often, sometimes, rarely or never. The first compilation phase was broken down into the following techniques: advanced search of databases (Excel, CSV, spreadsheets of Google Drive) on the internet, collection of databases through sources, web extraction of databases (scraping), compilation through contest or open collaboration projects (crowdsourcing), and creation of their own databases. The second phase of cleaning up was represented in the database edition task, while the context phase was arranged into database analysis. The fourth stage was focused on combining data; that is, the crossing of variables or databases. The fifth and final phase of the inverted pyramid of data journalism presented the greatest variety of options to fulfill the function of communicating. The seven techniques broken down were: creation of static infographics; creation of interactive visualizations; creation of digital narratives for multimedia platforms; publication of journalistic contents from social data (user profiles in social media and trends in the use of digital platforms); stories with interviews of human interest representative of the information contained in the databases; writing of news and customized journalistic contents from databases (age, sex, income, educational level, postal code) and development of tools and informative web applications built on database contents (Figure 2).

Additionally, the questionnaire included inquiries about the participants' country of origin, the media where they currently work, and their training in this specialization, specifically, if it is a result of their own initiative or driven by their companies. The questionnaire was developed in Qualtrics and Google Forms platforms and later distributed via email, WhatsApp, and Facebook to 44 journalists in Puerto Rico, 18 in Cuba and 13 in Dominican Republic. The analysis of the collected data was performed using pivot tables in Excel.

The analysis was segregated by countries and developed through a non-probabilistic convenience sample of 35 volunteers: 22 journalists from Puerto Rico[3] (50% of those addressed), seven from Cuba (39%), and six

3. Out of the 22 participants, 18 indicated the media where they work, while 4 did not answer the question.

Figure 2. How often do you perform the following tasks as part of your job?

Technique	Always	Often	Sometimes	Rarely	Never
1. Advanced search of datasets (Excel, CSV, Google Spreadsheet) in Internet					
2. Dataset supplied by a source					
3. Scraping databases online					
4. Create datasets yourself					
5. Edit datasets					
6. Analyze datasets					
7. Combine datasets or variables					
8. Preparing statics infographs (jpg, png)					
9. Preparing Interactive visualizations					
10. Create digital narrative for multimedia					
11. Compile data via contests or open collaboration projects (crowdsourcing)					
12. Create content based on social data (user profiles in social networks and trends in the use of digital platforms)					
13. Human interest story based on dataset					
14. Customize content based on dataset (age, sex, income, education level, zip code)					
15. Create web application or tool					

Source: prepared by the author based on the Inverted Pyramid of Data Journalism of Paul Bradshaw.

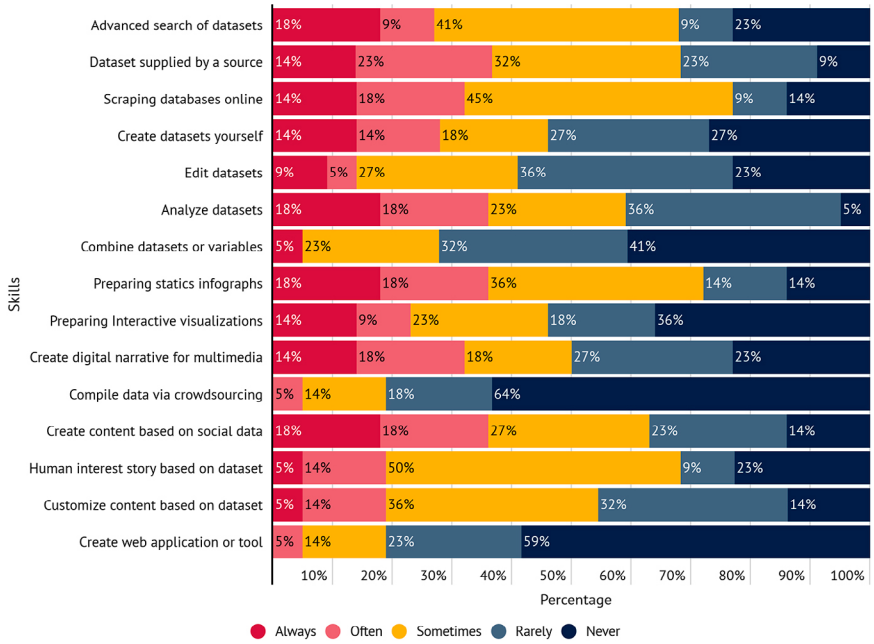
from the Dominican Republic (46%). The 18 journalists from Puerto Rico who indicated the media where they work are assigned to the following digital newsrooms: Metro PR (4), Centro de Periodismo Investigativo (3), El Nuevo Día (3), ONCE (3), Noticel (2), Todas PR (2), and Primera Hora (1). In the case of Cuban participants, they work in the media companies Postdata.club (2), Agencia de Noticias Cubana (1), Agencia EFE (1), Diario de Cuba (1), Fondo para las Naciones Unidas en Cuba (1), and Periodismo de Barrio (1). Likewise, respondents from the Dominican Republic work in the digital newsrooms of Diario Libre (2), Acento (1), Tu Sol Caribe (1), Actividades Artísticas.com (1), and Fiestas y Personalidades.com (1).

5. Results

A. Puerto Rico

The data journalism techniques most used by digital media journalists in Puerto Rico are the collection of databases through sources, the analysis of databases, the development of static infographics, and the publication of journalistic content based on data from social media (user profiles in social media and trends in the use of digital platforms). Each of these items was selected by 8 of the 22 participants (36%) as a skill that they "always" or "often" use in their work.

Figure 3. Frequency of use of data journalism techniques in Puerto Rico (2019).



Source: prepared by the author based on the answers provided by the 22 respondents.

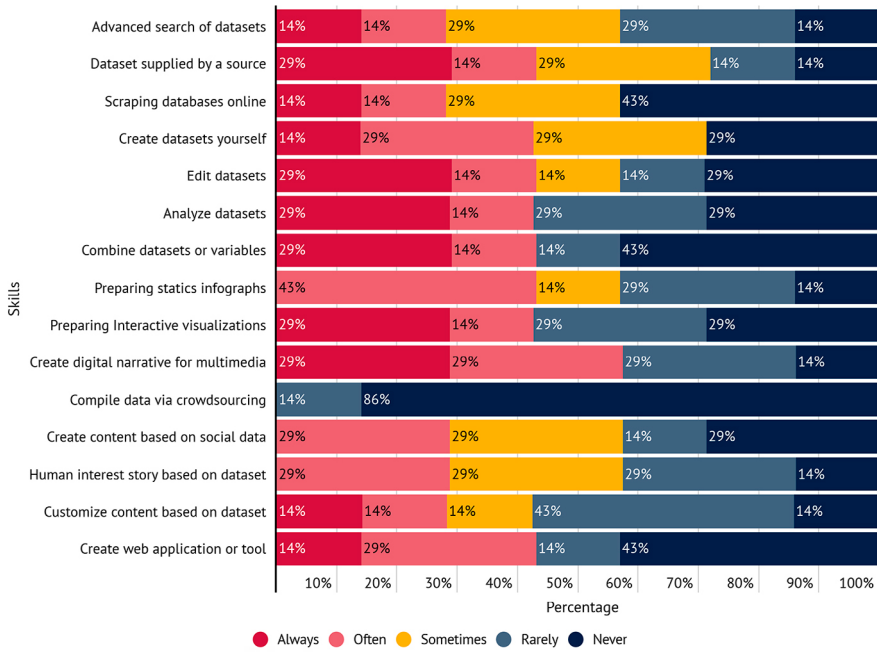
Furthermore, the collection of data through contests or open collaboration projects (crowdsourcing) and the development of informative web tools and applications based on database contents were the actions that respondents identified as the least used. In each of these two modalities, 18 participants answered that they "never" or "rarely" (82%) use them. The third data journalism skill least used by journalists in Puerto Rico is the crossing of variables or databases, as declared by 16 of the respondents (73%).

Regarding the training received on data journalism, 18 of the 22 professionals (82%) indicated that they have received some type of preparation. Of those 18 respondents, 12 participated in a workshop or professional practical seminar (67%), four attended a university class (22%), and two (11%) a Massive Open Online Course (MOOC). Likewise, six (33%) acknowledged that the professional training received stemmed from the company's initiative; conversely, 12 of the participants (67%) undertook this training on their own initiative.

B. Cuba

Cuban digital journalists who participated in the study identified the work of creating a digital narrative for multimedia platforms as the most used in their

Figure 4. Frequency of use of data journalism techniques in Cuba (2019)



Source: prepared by the author based on the answers provided by the 7 respondents.

professional performance. Four of the participants (57%) indicated that they do it "always" or "often." Other techniques frequently used by respondents are collecting databases through a source, creating them by themselves, editing, analyzing, combining databases, preparing static infographics and interactive visualizations, as well as developing tools or web applications. Each of these skills was selected by three participants (43%) as "always" or "often" used.

Moreover, the technique that the journalists do not use is the collection of data through contests or open collaboration projects (crowdsourcing). This option, marked by 100% of the seven participants in the "never" or "rarely" options, is followed by the development of tools and informative web applications based on database contents, cross-referencing of variables or databases, preparing interactive visualizations, customizing contents from databases, and analyzing databases. In all these cases, the percentage of professionals who responded that they use them "never" or "rarely" was 57%.

When analyzing the answers related to training in data journalism, four of the participants (57%) indicated that they had received some type of preparation in this specialization, while three (43%) admitted to not having

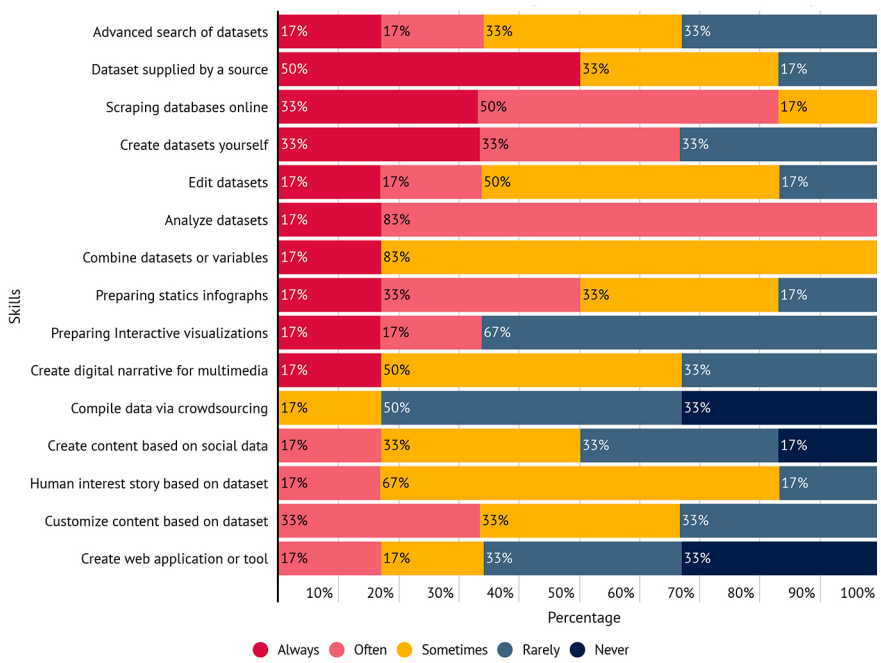
training in the area. Three out of the four journalists who studied the specialization in data journalism participated in workshops or seminars, and two attended on their own initiative.

C. Dominican Republic

The responses of the participants in the Dominican Republic suggested that the six digital journalists surveyed (100%) "always" or "often" analyze databases as part of their professional performance. Similarly, five (83%) said that they extract databases from the web, a method known as scraping, with the same frequency. The creation of their own databases is the third most used data journalism technique in digital media, as four professionals (67%) said that they "always" or "often" use this practice.

The collection of data through contests or open collaboration projects (crowdsourcing) is, on the contrary, the action least performed by this group. Five of the respondents (83%) stated that they "never" or "rarely" use it. This is followed by the creation of informative web tools and applications based on database contents and the preparation of interactive visualizations, with four respondents each (67%).

Figure 5. Frequency of use of data journalism techniques in Dominican Republic (2019)



Source: prepared by the author based on the answers provided by the 6 respondents.

All participants from the Dominican Republic claimed to have taken some form of data journalism training, mainly through a professional workshop or seminar (83%). Only one journalist received his training through a university course. Five of the journalists (83%) carried out the training on their own initiative, while the one remaining (17%) indicated that the initiative came from the company where he works.

6. Analysis and discussion

The processes that Bradshaw establishes in his inverted pyramid of data journalism vary in the frequency of use among journalists in Puerto Rico, Cuba, and the Dominican Republic, where the common denominator is the analysis of databases. This is the only skill that is identified among those used "always" or "often" in the three Spanish-speaking countries. Puerto Rico and Cuba also coincide in the compilation of databases through sources and in the creation of static infographics as techniques that are very useful in digital media. Supposing that the creation of static infographics is due to the absence of a programmer in the journalistic team (Crucianelli, 2013), a web designer in the writing or the lack of management of programming languages by reporters to create web visualizations, this research recommends digital journalists to use automated visualization tools available in freemium versions such as Datawrapper, Infogram or Flourish. In doing so, they can incorporate interactive graphics and visualizations without the need for prior programming knowledge or a larger budget.

Meanwhile, Cuban and Dominican journalists create their own databases. This may be due to the lack of official database inventories in both countries with which to write stories based on data journalism, as the government of the Dominican Republic launched its open data portal in 2015, while Cuba does not yet have a transparency and open data law or a web page that facilitates access to these (Hidalgo-Delgado, Mariño-Molerio, Amoroso-Fernández and Leiva-Mederos, 2018). Given this possible scenario, other useful strategies to work with databases are collaborations between media, journalists, activists, and technologists, as previously suggested by Kayser-Bril, Valeeva and Radchenko (2016).

Although just over half have followed some training related to data journalism, the Cuban respondents exhibit a greater variety in the types of tasks that they "always" or "often" apply when exercising data journalism, with up to nine different tasks from the 15 that are derived from the inverted pyramid. The techniques used include the five processes that Bradshaw establishes, with a greater emphasis on the communicating phase. These data indicate that Cuba's respondents work in digital media specialized in data journalism and/or that the preparation they received reaches a more advanced level, which allows them to put into practice all the techniques proposed, unlike Puerto Rico and the Dominican Republic.

Puerto Rico reveals less diversity of processes used frequently. The four techniques that digital journalists use correspond to the stages of compiling,

contextualizing, and communicating. However, they do not regularly incorporate tasks related to the cleaning and database combination phase. These two points are fundamental for the editing and handling of validated data, in addition to the possibility of enriching the data analysis by crossing variables or databases. These results show a contradiction between the high frequency of the analysis of databases and the low frequency of crossing variables and combination of databases among journalists in Puerto Rico. Among the techniques used to analyze databases is the crossing of variables and other skills such as the calculation of measures of statistical central tendency to determine the arithmetic mean, median, and mode. Therefore, if journalists assert that they "always" or "almost always" carry out data analysis, yet do not carry out cross-tabulation of variables or merge datasets with the same frequency, this indicates that the analyses that journalists perform stick to basic skills. They have not necessarily learned, mastered, or made the most of other functionalities available in data analysis software, such as pivot tables. The low frequency in the database combination technique may also be related to the difficulty involved in compiling databases through sources, the low availability of updated data inventories in the open data portals of that country and the recent approval of laws related to transparency and regulation of access to open data (Alvarado, 2019).

In the case of the Dominican Republic, respondents reply "always" or "often" to their use of three techniques. These belong to the stages of compiling and contextualizing, so they lack the usual performance of tasks related to the cleaning, combination, and communication of data. These three elements, part of the five that form the inverted pyramid of data journalism, are essential for the check, use, and in-depth analysis of databases, as well as the ability to achieve greater interaction with the audience through graphic presentations such as visualizations, infographics, narratives or stories of human interest, among other strategies, that encompass the fifth and final phase of communication.

The limited variety of strategies that digital journalists in the Dominican Republic use "always" or "often" contrasts with the 100% of the respondents who answered that they had followed some type of training in data journalism. Meanwhile, 18 out of the 22 professionals in Puerto Rico trained in this specialization, and four of them did so through a university course. Journalists in this country exhibit the highest number of people trained in this modality since only one other person in the Dominican Republic indicated the same option. Despite the high percentages of digital journalists who claimed to have received some training in the field of data journalism, neither of these two countries applies all the phases of the inverted pyramid proposed by Bradshaw. Conversely, in Cuba, the respondents demonstrated a greater diversity of strategies, while the percentage of journalists trained in this specialization is only a little more than half.

The reduced support that digital journalists receive from the journalistic companies in which they work to undergo training in new digital skills such

as data journalism is remarkable. While in Puerto Rico a third of the sample indicated that they had followed training at the initiative of the company, in the Dominican Republic there was only one person, and in Cuba, none.

7. Conclusions

The journalists from the three Caribbean countries agreed that the least used data journalism techniques are the collection of data through contests or open collaboration projects (crowdsourcing) and the development of tools and informative web applications based on database contents. Bradshaw argues that "crowdsourcing initiatives aimed at collecting data can also provide a social dimension to data," and highlights its great utility for collecting information "when it is difficult to obtain data in another way (Bradshaw, 2011a)." Therefore, this research recommends that digital media experiment with this type of strategy, which in turn establishes an additional connection with readers by integrating them in the data compilation process. Likewise, Bradshaw recognizes that the creation of informative web tools and applications is "the most complex way to communicate the results of data journalism." However, he acknowledges that its degree of complexity also translates into commercial opportunities due to the barrier they also present for competitors. Following his argument, the findings of this research proposes that if Caribbean media are interested in undertaking this type of project, they should consider offering hyperlocal stories, since, according to Bradshaw, "the closer your information is to their action, the more valuable it is to the user (Bradshaw 2011b)."

Journalists in Puerto Rico and Cuba concur in a third skill that they rarely apply: the crossing of variables or databases. Emphasis is placed again on the importance of combining data to deepen their analysis and extract relevant insights that allow innovative stories to be produced. Moreover, professionals from Cuba and the Dominican Republic agree that they "never" or "rarely" produce interactive visualizations, which is a very effective strategy to promote the interactivity of the contents in the increasingly used mobile support. The use of web tools that do not require programming knowledge such as Infogram, Datawrapper, and Flourish is recommended for the automated creation of interactive visualizations.

The results related to the strategies used in the three countries show that reporters do not use several stages of the inverted pyramid of data journalism. This may be due to the fact that they assume an advanced knowledge of the specialty that they have not yet acquired, since most indicated that training has been through a professional workshop or practical seminar. These circumstances are an excellent opportunity for Communication faculties of the Latin American and Caribbean universities to broaden their academic offerings with specialized courses in data journalism and visualizations as is already happening in American and European universities (Chaparro, 2014). This kind of specialized training will provide their students with the necessary

tools to practice data journalism in the media and offer innovative journalistic contents.

When analyzing the circumstances under which digital journalists followed some type of preparation, it is revealing that, in all countries, the training was mostly on their own initiative. The findings of this research demonstrated the need for digital media to promote the training of its professionals in specializations such as data journalism in order to develop an informative offer of added value through the publication of rigorous and interactive content.

The main limitation of this investigation was establishing contact with journalists from the Dominican Republic and Cuba and the low response to the request and follow-up efforts to participate in the study. This resulted in this initial study having a small sample in both countries. Another limitation is that the Qualtrics tool blocks access to users connected through a server in Cuba, which caused the questionnaire to be duplicated on the Google Forms platform so that journalists from that country could complete it.

This exploratory study opens up a variety of possibilities for future research, such as conducting a second descriptive phase that broadens the initial sample so that the results are more representative. It also represents an opportunity to expand it to other Latin American countries and include an analysis of the type of journalistic publications transmitted by digital media where data journalism techniques are incorporated. In doing so, we will be able to corroborate if the final journalistic product of these media reflects distinctive elements of this specialization, as well as the innovative nature that characterizes data journalism and has made it one of the main ways to offer added value contained in digital media of international importance.

Bibliographical references

ALVARADO, G. (November 22, 2019). "Gremios periodísticos piden cambios a las leyes de transparencia y datos abiertos". *El Nuevo Día.* [Viewed 11-22-2019]. Available from: <https://www.elnuevodia.com/noticias/locales/nota/gremiosperiodisticospidencambiosalasleyesdetransparenciaydatosabiertos-2531055/>

ANTÓN, S. (December 4, 2018). "Comenzará el 6 de diciembre servicio de internet en los móviles". *Granma*. [Viewed 6-23-2019]. Available from: <http://www.granma.cu/cuba/2018-12-04/comenzara-el-6-de-diciembre-servicio-de-internet-en-los-moviles-04-12-2018-20-12-27>

BENÍTEZ ROJO, A. (1989). *La isla que se repite: El Caribe y la perspectiva posmoderna*. Hanover, NH: Ediciones del Norte.

BRADSHAW, P. (2011a). "The inverted pyramid of data journalism – In Spanish". *Online Journalism Blog*. [Viewed 6-23-2019]. Available from: <https://onlinejournalismblog.com/2011/07/08/the-inverted-pyramid-of-data-journalism-in-spanish/>

— (2011b). "In Spanish: The inverted pyramid of data journalism part 2". *Online Journalism Blog*. [Viewed 6-23-2019]. Available from: <https://onlinejournalismblog.com/2011/07/14/in-spanish-the-inverted-pyramid-of-data-journalism-part-2/>

— (2013). *Data Journalism Heist: How to get in, get the data, and get the story out – and make sure nobody gets hurt!* UK: Leanpub.

BRADSHAW, P. and MASEDA, B. (2015). *Excel para periodistas: Cómo encontrar noticias en las hojas de cálculo*. Available from: <https://leanpub.com/excelparaperiodistas>.

BRADSHAW, P. and ROHUMAA, L. (2017). *Online journalism handbook* (2nd.ed.). Canada: Victoria. <https://doi.org/10.4324/9781315761428>

CASARES, A. and CHIVITE, L. (2015). "Laboratorio de Innovación Periodística: Contribución de la 'indisciplina' y de la 'multidisciplinariedad' ante la crisis de los medios". In: NOS ALDÁS, E.; ARÉVALO SALINAS, A. and FARNÉ, A. (Eds.) (2015). *#comunicambio: Comunicación y Sociedad Civil para el Cambio Social*. Madrid: Fragua, 386-398.

CHAPARRO DOMÍNGUEZ, M. (2014). "Nuevas formas informativas: el periodismo de datos y su enseñanza en el contexto universitario". *Historia y Comunicación Social*, 19, 43-54. <https://doi.org/10.5209/rev_HICS.2014.v19.45009>

— (2013). "La evolución del periodismo de precisión: el blog de *The Guardian* sobre periodismo de datos". In: F. SEGADO BOJ (Presidencia). *I Congreso Internacional de Comunicación y Sociedad Digital*.

COLÓN, E. (2017). *Media structures and the Press in Cuba, Dominican Republic and Puerto Rico, the Repeating Islands*. <https://doi.org/10.13140/RG.2.2.27170.02241>

CRUCIANELLI, S. (2013). *Herramientas digitales para periodistas*. Austin: Knight Center for Journalism in the Americas. [Viewed 5-13-2019]. Available from: <https://journalismcourses.org/wp-content/uploads/2020/08/Digital-Tools-for-Journalists-esp.pdf>

DARIAS, I. and SUÁREZ, M. (2016). "Cuba". In: SALAVERRÍA, R. (Ed.). *Ciberperiodismo en Iberoamérica*. Madrid: Fundación Telefónica and Editorial Ariel.

Data Journalism (Interdisciplinary). Missouri School of Journalism, University of Missouri, 2019. [Viewed 6-26-2019]. Available from: <https://journalism.missouri.edu/programs/data-journalism-interdisciplinary/>

DUNWOODY, S. and GRIFFIN, R.J. (2013). "Statistical Reasoning in Journalism Education". *Science Communication*, 35 (4), 528-538. <https://doi.org/10.1177/1075547012475227>

EGAWHARY, E. and O'MURCHU, C. (2012). "Data Journalism (CAR)". *The Center for Investigative Journalism*. [Viewed 6-26-2019]. Available from: <http://issuu.com/tcij/docs/data_journalism_book?e=2989993/2615602>

GONZÁLEZ, L. (2016). "Puerto Rico". In: SALAVERRÍA, R. (Ed.). *Ciberperiodismo en Iberoamérica*. Madrid: Fundación Telefónica and Editorial Ariel.

GRAY, J.; BOUNEGRU, L. and CHAMBERS, L. (Eds.) (2012). *Manual de periodismo de datos*. Sebastopol, CA: O'Reilly.

HIDALGO-DELGADO, Y.; MARIÑO-MOLERIO, A.; AMOROSO-FERNÁNDEZ, Y. and LEIVA-MEDERO, A. (2018). "Algunas reflexiones sobre los datos abiertos enlazados en Cuba". *Revista Cubana de Información en Ciencias de la Salud*, 29 (4). Available from: <http://www.acimed.sld.cu/index.php/acimed/article/view/1271/793>.

HOLOVATY, A. (2006). "A fundamental way newspaper sites need to change". [Viewed 11-22-2019]. Available from: <http://www.holovaty.com/writing/fundamental-change/>

HUDSON, M. (2017, April 7). "Panama Paper Wins Pulitzer Prize". *International Consortium of Investigative Journalists*. [Viewed 6-23-2019]. Available from: <https://www.icij.org/blog/2017/04/panama-papers-wins-pulitzer-prize/>

Investigative Reporting Specialization. Philip Merrill College of Journalism, University of Maryland, 2019. [Viewed 6-26-2019]. Available from: <https://merrill.umd.edu/academics/undergraduate-programs/specializations/>

Journalism B.A. School of Journalism & Mass Communication, University of Minnesota, 2005. [Viewed 6-26-2019]. Available from: <https://onestop2.umn.edu/pcas/viewCatalogProgram.do?programID=2271&strm=&campus=UMNTC>

Journalism Degree Requirements. American University, 2019. [Viewed 6-26-2019]. Available from: <https://www.american.edu/soc/journalism/BA/BA-JOURNO-REQUIREMENTS.cfm>

KAPLAN, D. (2013). "Global Investigative Journalism: Strategies for Support". [Viewed 6-23-2019]. Available from: <http://www.cima.ned.org/publication/global-investigative-journalism-strategies-for-support/>

KAYSER-BRIL, N.; VALEEVA, A. and RADCHENKO, I. (2016). "Transformation of Communication Processes: Data Journalism". Available from: <https://www.researchgate.net/publication/302381355_Transformation_of_Communication_Processes_Data_Journalism>

Ley Núm. 122 del año 2019: Ley de Datos Abiertos del Gobierno de Puerto Rico. Lexjuris Puerto Rico. August, 1, 2019 [Viewed 11-24-2019]. Available from: <http://www.lexjuris.com/lexlex/Leyes2019/lexl2019122.htm>

Ley Núm. 141 del año 2019: Ley de Transparencia y Procedimiento Expedito para el Acceso a la Información Pública. Lexjuris Puerto Rico. August 1, 2019 [Viewed 11-24-2019]. Available from: <http://www.lexjuris.com/lexlex/Leyes2019/lexl2019141.htm>

LÓPEZ HIDALGO, A. and UFARTE RUIZ, M. J. (2016). "Laboratorios de periodismo en España: Nuevas narrativas y retos de futuro". *Ámbitos*, 34, 1-12. Available from: <http://hdl.handle.net/11441/60784>

LORENZ, M. (2016). "Why Journalists Should Use Data. Data Journalism Handbook". [Viewed 6-25-2019]. Available from: <http://datajournalismhandbook.org/1.0/en/introduction_1.html>

MEYER, P. (2012). *Paper Route: Finding My Way to Precision Journalism*. Bloomington, IN: iUniverse.

ORTIZ, I. (2016). "República Dominicana". In: SALAVERRÍA, R. (Ed.). *Ciberperiodismo en Iberoamérica*. Madrid: Fundación Telefónica and Editorial Ariel.

ROGERS, S. (2014). "Why data reporters should ask themselves the five Ws of journalism". [Viewed 6-25-2019]. Available from: <https://ijnet.org/en/story/why-data-reporters-should-ask-themselves-five-ws-journalism>

— (2016). "Data journalism matters more now than ever before". [Viewed 6-25-2019]. Available from: <https://simonrogers.net/2016/03/07/data-journalism-matters-more-now-than-ever-before/>

SÁDABA, C. and SALAVERRÍA, R. (2016). "La innovación y los cibermedios: los labs". In: SÁDABA, C.; GARCÍA, J.A. and MARTÍNEZ-COSTA, M. (Coords.). *Innovación y desarrollo de los cibermedios en España*. Navarre, Spain: Ediciones Universidad de Navarra, 41-47.

SALAVERRÍA, R. (2015). "Los labs como fórmula de innovación en los medios". *El Profesional de la Información*, 24 (4), Nuevos medios II, July-August, 397-404. <https://doi.org/10.3145/epi.2015.jul.06>

— (2016). *Ciberperiodismo en Iberoamérica*. Madrid: Fundación Telefónica and Editorial Ariel.

SALAVERRÍA, R.; MARTÍNEZ-COSTA, P. and BREINER, J. (2018). "Mapa de los cibermedios de España en 2018: análisis cuantitativo". *Revista Latina de Comunicación Social*, 73, 1034-1053. <https://doi.org/10.4185/RLCS-2018-1295>

SALAVERRÍA, R.; RIVERA-ROGEL, D. and GONZÁLEZ-CÓRDOVA, M.P. (2019). "Taxonomía del periodismo digital en Iberoamérica: evolución en las dos décadas digitales". In: RIVERA-ROGEL, D. and ROMERO-RODRÍGUEZ, L.M. (Eds.). *Comunicación digital: Actualidad, retos y prospectivas*. Lima: Pearson Educación de Perú, S.A, 263-285.

Specialization in Digital Journalism (Bachelor of Arts in Media). Indiana University Bloomington, 2019. [Viewed 6-26-2019]. Available from: <https://bulletin.college.indiana.edu/programs/index.html?program=mdas06>

SURLIN, S.H. and SODERLUND, W.C. (1990). *Mass Media and the Caribbean*. New York: Gordon and Breach.

The 2017 Pulitzer Prize Winner in Explanatory Reporting. The Pulitzer Prizes, 2019. [Viewed 5-13-2019]. Available from: <https://www.pulitzer.org/winners/international-consortium-investigative-journalists-mcclatchy-and-miami-herald>

The 2016 Pulitzer Prize Winner in National Reporting. The Pulitzer Prizes, 2019. [Viewed 5-13-2019]. Available from: <https://www.pulitzer.org/winners/washington-post-staff>

The 2015 Pulitzer Prize Winner in Investigative Reporting. The Pulitzer Prizes, 2019. [Viewed 5-13-2019]. Available from: <https://www.pulitzer.org/winners/wall-street-journal-staff>

The 2013 Pulitzer Prize Winner in Public Service. The Pulitzer Prizes, 2019. [Viewed 5-13-2019]. Available from: <https://www.pulitzer.org/winners/sun-sentinel>

The 2012 Pulitzer Prize Winner in Investigative Reporting. The Pulitzer Prizes, 2019. [Viewed 5-13-2019]. Available from: <https://www.pulitzer.org/winners/michael-j-berens-and-ken-armstrong>

The 2011 Pulitzer Prize Winner in Investigative Reporting. The Pulitzer Prizes, 2019. [Viewed 5-13-2019]. Available from: <https://www.pulitzer.org/winners/paige-st-john>

TREADWELL, G.; ROSS, T.; LEE, A. and LOWENSTEIN, J. K. (2016). "A Numbers Game: Two Case Studies in Teaching Data Journalism". *Journalism & Mass Communication Educator*, 71 (3), 297-308. <https://doi.org/10.1177/1077695816665215>

VERBORGH, R. and DE WILDE, M. (2013). *Using OpenRefine*. Birmingham, UK: Packt Publishing.