COGITATIO

Article

Media Roles in the Online News Domain: Authorities and Emergent Audience Brokers

Sílvia Majó-Vázquez [1,*], Ana S. Cardenal [2], Oleguer Sagarra [3] and Pol Colomer de Simón [3]

[1] Reuters Institute for the Study of Journalism, University of Oxford, Oxford, OX2 6PS, UK;
E-Mail: silvia.majo-vazquez@politics.ox.ac.uk
[2] Department of Politics, Open University of Catalonia, 08860 Barcelona, Spain; E-Mail: acardenal@uoc.edu
[3] Dribia Data Research S. L., 08018 Barcelona, Spain; E-Mails: ula@dribia.com (O.S.), pol@dribia.com (P.C.d.S.)

* Corresponding author

Abstract
This article empirically tests the role of legacy and digital-born news media, mapping the patterns of audience navigation across news sources and the relationship between news providers. We borrow tools from network science to bring evidence that suggest legacy news media retain control of the most central positions in the online news domain. Great progress has been made in discussing theoretically the impact of the Internet on the news media ecology. Less research attention, however, has been given to empirically testing changes in the role of legacy media and the rising prominence of digital-born outlets. To fill this gap, in this study we use the hyperlink-induced topic search algorithm, which identifies authorities by means of a hyperlink network, to show that legacy media are still the most authoritative sources in the media ecology. To further substantiate their dominant role, we also examine the structural position of news providers in the audience network. We gather navigation data from a panel of 30,000 people and use it to reproduce the network of patterns of news consumption. While legacy news media retain control of the brokerage positions for the general population, our analysis—focused on patterns of young news consumers—reveals that new digital outlets also occupy relevant positions to control the audience flow. The results of this study have substantive implications for our understanding of news organizations' roles and how they attain authority in the digital age.

Issue
This article is part of the issue "Digital Native News Media: Trends and Challenges" edited by Ramón Salaverría (University of Navarra, Spain).

1. Introduction

Great progress has been made in discussing theoretically the impact of the Internet on the news media ecology. Less research attention, however, has been given to empirically testing changes in the role of legacy media and the rising prominence of digital-born outlets. The growth of the popularity of digital-born outlets as news providers (Nicholls, Shabbir, & Nielsen, 2016) and the increasing role of the Internet as a main access point for news consumption have motived scholars to con-

tend that there has been a reconfiguration of media power. As a result, digital-born media might be challenging the classic monopoly of legacy brands as authoritative information providers and audience builders. Hence, according to some, new elites might be emerging (Castells, 2009; Chadwick, 2013; Couldry & Curran, 2003; Gurevitch, Coleman, & Blumler, 2009; Hermida & Thurman, 2008; Jarvis, 2016; Pavlik, 2001). Yet, whether this is the case or not remains largely untested. There are a few empirical studies to support this line of work (Meraz, 2009), and they are contested by evidence sug-

gesting that media functions, i.e., agenda setting or gatekeeping, are still in the hands of legacy media (Coleman & McCombs, 2007; Lee, 2007; Park, Ko, Lee, & Song, 2013). Following a convention in the literature, in this study the term legacy media refers to those outlets that predate the Internet era and the term digital-born media identifies news providers that do not have an offline counterpart.

Legacy media are struggling to deal with the endless technological upheaval and to secure their positions as leading information providers in the online domain (Althaus & Tewksbury, 2000; Cagé, 2016; Downie & Schudson, 2009; Fenton, 2010; McChesney & Pickard, 2011). Their paper editions' readerships shrink and so do their revenues (cf. Anderson, Bell, & Shirky, 2012; Newman, Fletcher, Kalogeropoulos, & Nielsen, 2019). At the same time, digital-only competitors make it more difficult for legacy media across geographies to engage a new loyal public, especially when it comes to paying for content and attracting a young target audience (Cage, Viaud, & Herve, 2015; Pew Research Center, 2015; Suárez, 2020).

Latin America is one of the world regions where the birth and growth of digital-only media has been most dramatic profound (Salaverría, Sádaba, Breiner, & Warner, 2019). There, they are "shaking the market" with larger audience figures than those of legacy and well-established outlets (Harlow & Salaverría, 2016; Salaverría, 2016). In many of these countries, as well as in other parts of the world, digital-born outlets have also shown a faster adoption of innovation than their legacy counterparts. This ability has not only been recognized by legacy media but has also been highlighted in the meta-journalistic discourses that digital-born media have used to publicly present themselves as unique news sources in the online domain (Carlson & Usher, 2016).

In parallel to the rise of digital-born outlets, several examples speak to the efforts that legacy media are continuing to protect their role as the main information providers in the digital age (Castells, 2009; Chadwick, 2013; Cook, 1998; Garnham, 1995; Meraz & Papacharissi, 2013). In the UK, for instance, *The Independent* announced an unprecedented digital-only format movement ("The Independent becomes," 2016) and ceased its print edition. Similarly, one of the top Spanish news brands, *El País*, suggested the possibility of ending its paper edition in the short term (Caño, 2016). Finally, *The New York Times* and, more recently, the *Der Spiegel* commissioned innovation reports that contain detailed critical analyses regarding the organizations failing to successfully embrace the digital sphere ("Mitarbeiter fordern Revolution von unten," 2016; Wills, 2014). *The New York Times* acknowledged: "We have always cared about the reach and impact of our work, but we haven't done enough to crack that code in the digital era" (Wills, 2014, p. 3). Perhaps consistent with those concerns, the newspaper has recently looked closely for new talent from among the digital-born competitors (Franklin, 2018; Smith, 2020).

These examples illustrate the importance that legacy media give to the online domain. Nonetheless, citizens increasingly turn to the Internet—including social media—as their main access to political information, and in some countries they already prefer online sources for news rather than print editions or even television (Newman et al., 2019).

Underlying the actions of the legacy brands described above—all aimed at concentrating efforts on the web—there is a broad question related to the function that one of their main competitors, digital-born outlets, are playing in the provision of information. Certainly, in the current hybrid media system, there is a confluence of legacy and digital-born brands and a negotiation for the re-allocation of roles among them (Carlson, 2017; Chadwick, 2013). Yet, whether this implies a decline in the role of legacy media and a rise of digital-born outlets must be put to an empirical test.

This leads to the main question guiding the analysis that follows: Are legacy media the central actors in the online news domain? Their audience rates might prove they continue to be. Overall, in many regions in the world, legacy media still retain higher proportions of attention, measured in the number of unique visitors, compared with those of digital-born brands. However, we argue that the capacity for generating traffic is merely one factor that determines the relevance of a news source. To understand how and when news media attains authority in the online domain, it is necessary to identify those who hold central positions in the audience flow and are considered providers of unique content by digital audiences and by news providers themselves.

To assess media authority, research traditionally looked into collective narratives by journalists (Carlson, 2007; Zelizer, 1990). Later, Carlson (2017) noted that not a single variable can explain how journalists, and therefore news media, attain authority. In his seminal attempt to explain journalistic authority, he instead advocates for a relational perspective that considers an array of actors inside and outside the newsroom, e.g., reporters, sources, audiences, critics. In line with this theoretical work, we propose an observational approach to evaluate the role of legacy and digital-born organizations from a two-dimensional perspective. We look at the level of authority that news sources confer to each other and how audiences navigate them, thereby revealing their dominance of the online audience flow.

We assess the centrality of news sources by looking at the two main dimensions of the online news domain. First, we analyse the supply side by reproducing the network of hyperlink connections among news providers and measuring the authority of the news media. For this purpose, we draw on the methodological framework by Kleinberg (1999) and previous work by Weber and Monge (2011). More specifically, we borrow tools from network science to measure the centrality of news outlets in the flow of news information. We analyse the extent to which news sources quote others' content, which

is a proxy to identify the level of authority that one conferred to the others. This approach has largely been used in the field of information retrieval and by popular search engines, e.g., Google, to return meaningful results (Newman, 2010). Recently, it has been applied to fact-check news sources too (Kosslyn & Yu, 2017). Our goal here is bringing evidence to assess whether legacy media retain control of the most central positions in the online news domain by measuring whether they are endorsed as authorities, i.e., sources of reference by their peers.

Second, to further substantiate the centrality of the news sources, we assess their role on the demand side of the news domain and the flow of audience across them. We gathered online navigation data from a representative panel of 30,000 people in Spain and reproduce their patterns of news consumption drawing on previous work in the field of audience networks (Ksiazek, 2011; Webster & Ksiazek, 2012). In doing this, we test whether the most central sources in the provision of news content also hold central positions to control the audience flow. We offer this analysis in the Spanish context because it allows a comparative dimension to the few empirical analyses on the centrality of news media which are largely focused on the US (Meraz, 2009; Meraz & Papacharissi, 2013; Meyer, 2004; Weber & Monge, 2011).

Third, we review the previous work on the impact of the Internet in the news media ecology and how legacy media have retained attained authority while facing the emergence of digital-born outlets. We also offer a presentation of the Spanish media sector. Fourth, we introduce the data and methods and discuss their potential to tackle the relationships among different types of news organizations and measure media centrality in the news domain. Fifth, in the results section, we present the outcomes of this role identification analysis. Finally, we discuss our results and their implications for the theoretical accounts about how the Internet is reconfiguring the digital news domain.

2. News Organizations in the Digital Era

The patterns of interaction between legacy and digital-born media and the re-allocation of roles among them have lately spurred much interest. In the early years of the 21st century, Dutton, Gillett, McKnight, and Peltu (2004) already noted the disruptive impact of digital technologies on the reconfiguration of power among several types of actors and organizations. Media institutions have not been an exception. Some scholars have argued that the web has created new spaces for power (Bennett, 2003), new hierarchies have emerged in the online domain (Mansell, 2004) and, ultimately, the web has recast the roles of actors in the media ecology (Carlson, 2007; Gurevitch et al., 2009; Tandoc & Jenkins, 2017).

The web gives audiences greater access to a broader range of facts, data, and opinions and changes the traditional journalistic standards for the provision of news (Fortunati et al., 2009; Riordan, 2014). Communicating

what matters to the public is no longer a lineal process from sources to journalist to be finally transmitted to mass audiences. Instead, the web has created a new distinctive scenario where providing information and news is no longer the exclusive task of legacy organizations (Croteau, 2006). The mass-self-communication system (Castells, 2007) or the many-to-many model of information theoretically puts the smallest news providers on an equal footing with the transnational conglomerates (Rheingold, as cited in Fenton, 2010). Consequently, people do not only rely on traditional media to make sense of the myriad of information sources around them. Information control or gatekeeping, according to this strand of research, is now shared either with audiences that perform as content producers or with a wide range of new actors, some of whom act as journalists, providing newsworthy content too. We categorize them here under the label of digital-born outlets, e.g., blogs, news sites, aggregators or niche outlets.

Before all of these changes, legacy media took advantage of scarcity, exclusivity and control of information. They did so for much of the 20th century (Lewis, 2012, p. 311) to secure their positions as elites in the political communication process. They controlled information dissemination in a structured environment and shaped the social order which ultimately allowed them to operate with some authority (Robinson, 2007, p. 307).

Traditionally, to reify their authoritative status to audiences, journalists from legacy media have used narrative and placed themselves at the centre of stories (Zelizer, 1990). Through meta-journalistic discourse, they have also set boundaries and excluded other actors from playing a role in the news sector (Carlson, 2017). They have identified and credited those who have the authority to provide news information and those who do not; frequently overlooking the role of the audience in legitimizing that authority.

Digital technologies have, however, shaken up the traditional role of the legacy media and redefined who attains authority in the online domain. This has necessarily broadened the understanding of media authority. With the advent of the web, digital-born outlets emerged as new empowered actors (Singer, 2007). Some of them initially established themselves in opposition to the legacy vehicles for information production (Carlson, 2007). Certainly, they challenged the traditional role of legacy outlets and even the very notion of journalistic professionalism (Singer, 2003). In turn, legacy media received them with a combination of uncertainty (Tandoc & Jenkins, 2017) and resistance to legitimize them as authoritative sources.

Similar to legacy organizations, digital-born outlets can reach a mass audience. Although they have much smaller structures, even a blog by a single author can reach thousands of people today. Notably, the web has radically lowered the cost of distribution (Cage et al., 2015). Their reach and also their relevance, especially during major political events, are a proxy to understand

their ability to challenge the monopoly of media authority, which was once exclusively attained by legacy outlets (Majó-Vázquez, Zhao, Nielsen, 2017; Majó-Vázquez, Zhao, Nurse, & Nielsen, 2017).

Among all the changes brought about by the advent of the Internet in the media sector, Livingstone considers that potentially the most radical change is the shift from one-way mass communication towards more interactive communication (1999, p. 61). Along these lines, digital-born outlets have given an active role to audiences in the production of information. They have blurred the distinction between producers of news content and the audiences themselves (Croteau, 2006), a relationship well-established for decades by legacy media to consolidate their authority.

Digital-born outlets have engaged diverse actors in horizontal and conversational practices to filter and promote content, blending broadcasting with social conventions (Meraz & Papacharissi, 2013, p. 142). They have shared and transformed journalistic routines, once considered exclusive to legacy news media, like gatekeeping and agenda setting, with a broader set of actors and the audience. Digital-born outlets have diversified standard journalistic narratives, formats and rituals. They provide easily sharable content (Riordan, 2014), multimedia news products, and take advantage of new forms of collaboration, analytic tools and sources of data (Anderson et al., 2012).

Alternative sources of funding have also become a growing mechanism to differentiate new and legacy outlets and to interact and engage with the public. Beyond the classical sources of media funding, such as advertisements or even government subsidies (Brogi, Ginsborg, Ostling, Parcu, & Simunjak, 2015; Colino, 2013), digital-born outlets have also sought alternative sources of income to sustain their journalistic practices, highlighting their independence and increasing audience participation.

All in all, digital-born outlets have been agents of innovation (Carlson & Usher, 2016) and, therefore, have changed how journalism can be understood and how media authority in the news domain is attained. Interestingly, they have addressed audiences by means of metadiscourses where they create distance from legacy models of journalism and stress their ability to construct ideal-types of news thanks to the affordances of the web (Carlson & Usher, 2016). Yet again, whether all this has positioned them as central actors in the news domain and has allowed them to challenge the authoritative role of legacy media has not been empirically studied. We aim to fill this gap.

2.1. The Spanish News Media Landscape

To this end, we focus on the Spanish online news domain which has also been affected by the above mentioned changes. It is important to note that the Spanish media outlets are overtly partisan (Hallin & Mancini,

2004) and the media system greatly suffered from the impact of the economic crisis that hit the country from 2008 to 2012. Since then, legacy media have struggled to overcome both the media crisis and the economic crisis. Between 2008 and 2013, the unemployment rate for Spanish journalists rose 132%, the communications sector lost 11,151 jobs and 284 news media outlets shut down (Press Association of Madrid, 2013). However, on the upside of that adverse scenario, 458 new outlets were founded during that very same period. The vast majority of them were digital-born (Press Association of Madrid, 2015) and led by journalists who used to work for legacy organizations (Minder, 2015; Schoepp, 2016). Prominent examples are *Elespanol.com*, founded by the former director of the second largest legacy outlet in Spain, *El Mundo*, and *ctxt.es*, founded by former senior journalists from *El País*.

In sum, theoretical accounts of the reallocation of roles in the digital domain contend digital-born outlets have shown their ability to transition from volume to value, and that authority has been conferred to them by building relationships with people based on relevance (Jarvis, 2016). If that is the case, they should hold central positions in the news media domain. This leads to the following research question:

RQ1: To what extent are digital-born outlets central actors in the flow of news content by being recognized as authoritative sources of information by other news providers?

Yet, as we have argued, we propose a multidimensional approach to the centrality of the news media. Therefore, we assess whether authoritative sources of information also have central positions in the control of audience flow. Central sources of information should not only be regarded as unique as authorities by their peers but they should also have a certain potential control over the way audiences navigate the news domain. This idea is illustrated in network theory by measuring betweenness centrality of nodes. Media outlets with higher betweenness centrality lie on the audience shortest path between other news sources and therefore may have considerable influence over the overall audience behaviour. This leads to our following sub-question:

RQ1.2: To what extent do legacy and digital-born outlets hold brokerage positions in the audience network?

Finally, we expect that the underlying structure of the audience network differs if we consider specific patterns of news consumption across demographics. Studies that look at the use of news by young generations unveil distinctive media diets. Some have pointed out that legacy status is less important in shaping their online navigation (Taneja, Wu, & Edgerly, 2018) and that they are more keen on using social networks and mobile devices

to keep up with current affairs (Newman et al., 2019). They are also less likely to remember news sources' names (Mitchell, Gottfried, Shearer, & Lu, 2017), which might imply that brand awareness, which is especially important for legacy outlets, is less effective in attracting and engaging with younger audiences. Further, digital-born outlets are setting up designated channels to reach young audiences, which, in turn, increases the competition against legacy brands. In light of this specific scenario, our final research question examines the role of news providers across different age groups:

RQ1.3: Do legacy and digital-born outlets differ in their brokerage positions when considering different age groups?

3. Data and Methods

3.1. News Network

In this study, we use two different techniques to analyse two types of networks: the hyperlink network, also called news network, and the audience network. In total, we study six networks—five audience networks representing news consumption patterns of different age groups and one hyperlink network.

The news network represents the total number of hyperlinks that news media send to each other and is weighted and directed. Here, nodes are news media outlets, and an edge between two nodes exists if there is at least one hyperlink sent from one outlet to an external source. Two media outlets are disconnected if none of them has cited content from the other. The weight of the edges equals the total number of hyperlinks among each pair of nodes.

We argue that a hyperlink is a conservative measure of the authority that news media confer to competitors. It has been shown that linking has a very high cost barrier in the media ecology. In line with the *walled-garden philosophy*, news media rarely link to competitors to avoid guiding audience attention outside its boundaries (Chung, Nam, & Stefanone, 2012; Larsson, 2013; Napoli, 2008). Therefore, following previous research in the field of information retrieval (Kleinberg, Kumar, Raghavan, Rajagopalan, & Tomkins, 1999; Weber & Monge, 2011), we assume that when they do send a hyperlink to another news outlet, it is to acknowledge the unique value of its content. Algorithms behind search engines also determine the level of authoritativeness of a site by, among other indicators, the number of inlinks received from peers specialized in pointing to sources of information (hub) (Newman, 2010). Additionally, outlinks have also been used in the journalistic sector as proxies for trusted sources of information and, consequently, to build measures of media trust (Bale, Walmsley, Wustermann, Ericson, & Barber, 2018).

To map the news network, we start by selecting the top 44 news outlets in Spain in February 2015. We used

Alexa rankings ("Top sites in Spain," 2015) for selecting them. To test the accuracy of the data, we compared the initial ordering with a list provided by ComScore, the official digital audience meter in Spain. The correlation of the two rankings was 0.90. To this list, we added top digital outlets in Spain using the ComScore ranking. We ended up with a final list of 100 news outlets that we input as seeds in commercial software for crawlings (Voson). The web crawl was supervised to ensure that we extracted the maximum number of incoming and outgoing links from the deepest level. The process not only involved snowballing from seed sites to get outgoing hyperlinks but also querying a search engine API (Bing) to collect incoming hyperlinks. Our approach is consistent with previous research that shows that selecting seed sites according to well-established prominence (in our case audience reach) minimizes the bias of rendering peripheral sites as central (Weber & Monge, 2011). After page-grouping subdomains with their parent sites and pruning those nodes that were not news providers, the final network has 100 nodes and almost 800 edges or connections. Table 1 summarizes its main descriptive statistics (see Supplementary File for further descriptive analyses).

3.2. Audience Networks

Finally, we built our four audience networks in November 2015. We drew on the long history of audience duplication data and more recently online audience duplication data to define the ties in these networks as the shared audience among news media outlets and the nodes as news sources (Cooper, 1996; Goodhardt & Ehrenberg, 1969; Goodhardt, Ehrenberg, & Collins, 1987; Ksiazek, 2011; Webster & Ksiazek, 2012). Thus, in the four audience networks, ties stand for the total audience overlap between media *i* and media *j*. These networks are directed and weighted. Previous work in this field has also leverage on similar data to calculate the weight of the ties of audience networks as the amount of audience shared by news sources (Taneja & Webster, 2016). Notably, in this study we keep the direction of the ties too. This offers more granular information about audience behaviour (see Ksiazek, 2011, for further discussion on directed and undirected ties in audience networks). Because our data collection process for mapping the hyperlink and the audience networks had to take place at different points in time, we tested the volatility of the audience data by comparing media reach across a one-year period. According to the resulting correlations—the lowest score equals 0.90—we can assume fairly stable audience behaviour during the period of time that our study spans, with no major changes in the patterns of news consumption between February 2015 and November 2015.

We gathered the audience data from ComScore (Interactive Adversiting Bureau, 2011). This is a private provider which tracks the browsing behaviour of a representative panel of 30,000 Spanish people. To check the accuracy of the measurements, it combines observations

Table 1. Descriptive statistics of news and audience networks.

Statistic	News network Feb-15		Audience networks Nov-15	
	Full	News network	Total audience	Young audience
Size	16,659	100	108	93
Edges	32,628	799	10,387	8464
Density	0.000	0.081	0.890	0.989
Reciprocity	0.046	0.471	0.980	0.989
Clustering	0.006	0.367	0.930	1.000
Max. Inlinks	1223	32	38	92
Min. Inlinks	0	0	106	91
Max. Outlinks	704	49	0	92
Min. Outlinks	0	0	106	0
Max. Instrength	2793	1173	44150	5199
Min. Instrength	0	0	0	0
Max. Outstrength	1282	447	42120	2946
Min. Outstrength	0	0	0	0
Assortativity	−0.615	−0.166	−0.11	0

coming from the panel with other indicators, drawn from direct audience metrics from news media collected by a separate agency (Asociación para la Investigación de Medios de Comunicación, 2014). The audience networks represent observed data instead of reported news media diets, which have proven to be overstated and result in less accurate analyses (Guess, 2015; Prior, 2009).

Our audience provider only offers audience overlapping data for those sites that have been visited by a minimum of 16 panelists in a given month. As a consequence, the total number of nodes in the four networks varies. We reproduced the online audience navigation patterns for the general population and people aged between 18–24, 25–34, 35–54 and more than 55 years old.

3.3. Methods

Our main goal is to identify the role of news providers in the digital domain. We define as authoritative sources of information those media outlets that are frequently linked by their peers and by doing so we can provide evidence to identify sources of information that potentially control the flow of news. To this end, we use two different techniques borrowed from network science. We first measure the level of authority of news media using the hyperlink-induced topic search algorithm developed by Kleinberg (1999). Following previous work in the area of Link Analysis Rank algorithm, we assume that a hyperlink from node i to node j denotes an endorsement for the quality of the page j (Borodin, Roberts, Rosenthal, & Tsaparas, 2005; Kleinberg et al., 1999; Weber & Monge, 2011) and that the authority score is a proxy to identify the most central news sources online.

The hyperlink-induced topic search algorithm provides a two-level propagation scheme, where endorsement is conferred on authorities through hubs. In other words, every page has two identities: The hub identity

captures the quality of the page as a pointer of useful news sources, and the authority identity, our main concept of interest, captures the quality of the page as a resource itself (Borodin et al., 2005, p. 235). We identify news providers' roles as authorities and hubs. The former are news media, which provide content highly cited by their peers, and the latter are news media that cite the worthiest sources of news information. Formally, the authority centrality of a vertex is defined to be proportional to the sum of the hubs centralities of the vertices that point to it:

$$x_i = \alpha \sum_j A_{ij} y_j,$$

where A_{ij} is an element of the weighted adjacency matrix, y_j stands for the hub centrality and α is a constant. Consistently, the hub centrality of a vertex is proportional to the sum of the authority centralities of the vertices it points to:

$$y_i = \beta \sum_j A_{ji} x_j,$$

where β is also a constant (for more information see Kleinberg, 1999; Newman, 2010; Weber & Monge, 2011). Additionally, we use a random benchmark known as the 'configuration multi-edge model' in the network literature (Sagarra, Font-Clos, Pérez-Vicente, & Díaz-Guilera, 2014). It is specifically designed for weighted networks and matches the strength distribution of our news and audience networks. This benchmark aims to assess the significance of the authority and hub scores previously calculated and its departure from what we could expect if the hyperlinks in the news network were sent randomly. The null model is based on the reconstruction of 1,000 random networks with exactly the same strength distribution as our observed news network and maximally random in all other respects. Because our news

network is directed, the null model is defined separately for incoming and outgoing ties (for more information see Sagarra et al., 2014).

We operationalize the role of the news sources based on multidimensional definition. Therefore, along with the authority analysis, we also assessed the potential control of news media over the audience flow. To this end, we measured the betweenness centrality score of all the news sites in the four audience networks that represent patterns of news consumption across ages. This approach allows bringing evidence to the role of the news providers as audience brokers across demographics. Formally, the brokerage scores are roughly equal to the number of the shortest paths between others that pass through a node. More interestingly, though, the resulting ranking of brokers for each network will identify those news providers that have higher potential control over the audience flow in each age group and can thereby engage wider audiences.

Since Freeman (1979) conceptualized the concept of brokerage, in the sociological literature there is a significant amount of work devoted to the study of this term in social networks (Coleman, 1988; Gonzalez-Bailon

& Wang, 2013). We employ here the formalization of betweenness centrality by Opsahl, Agneessens, and Skvoretz (2010), which generalizes the shortest path calculations for weighted networks matching the nature of our data. In other words, we not only took into account the shortest paths between nodes in our audience network but also the amount of audience that flows through them. As a result, our measure not only accounts for those news sources that more frequently mediate visits to other news sources but also for those that receive more visits themselves. In order to determine whether the brokerage scores are statistically significant (not due to random fluctuations of the audience flow), we use again the configuration multi-edge model.

4. Results

The identification of the media authorities, or the most reputable sources of news content in the digital domain, yields evidence that legacy news outlets broadly control this strategic position at the individual level. Figure 1 shows the authority scores of the two types of news providers under study: legacy and digital-born
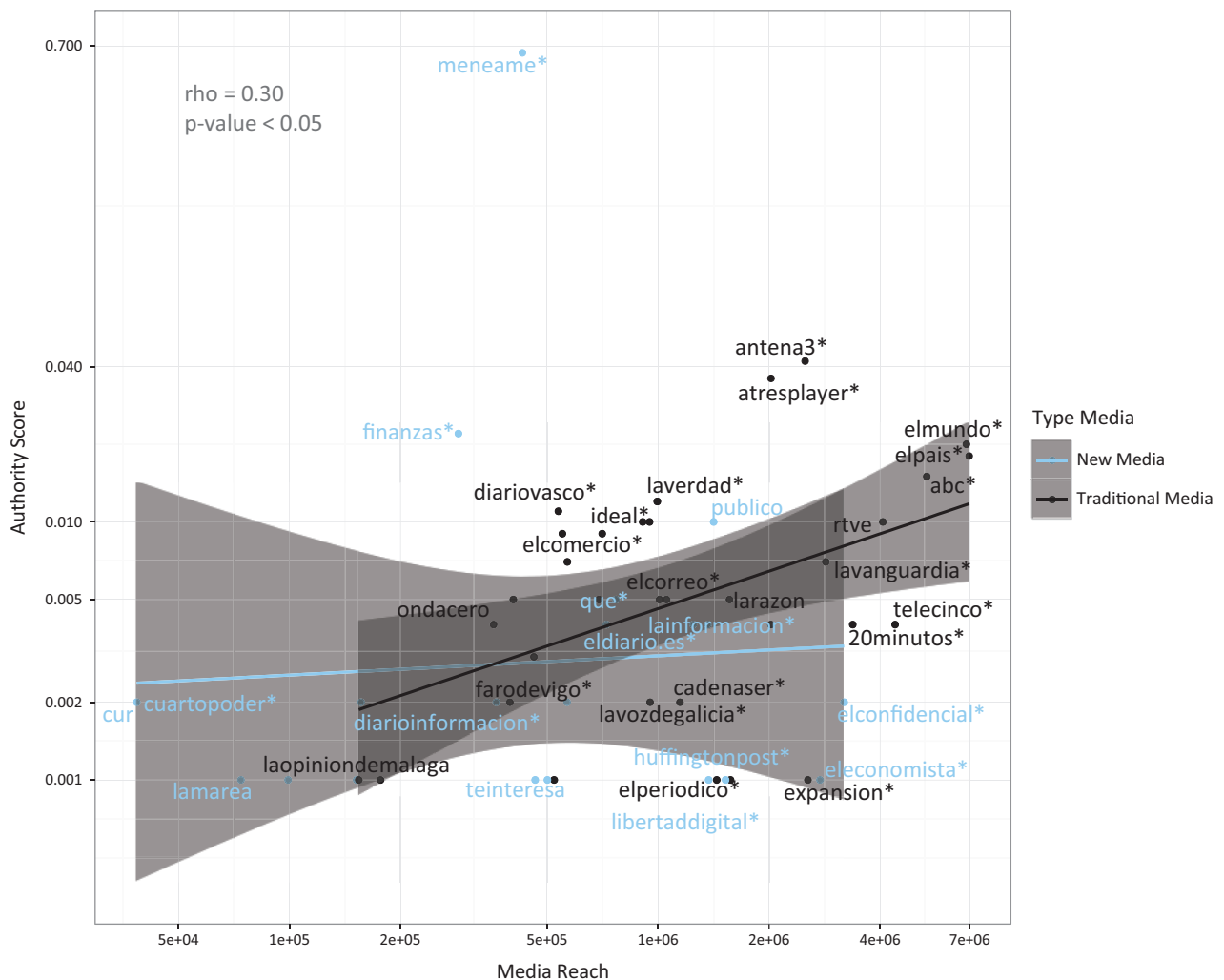

Figure 1. Association between authority scores in the news network and audience reach. Note: * significant score.

media. The distribution of the scores suggests that legacy brands are more frequently considered sources of valuable content, i.e., more authoritative than digital-born media. There are only two exceptions. Among the top 10 authorities identified in the hyperlink network, we found the aggregator *Meneame.es* and the business outlet *Finanzas.es*. One plausible explanation for the outstanding position of *Meneame.es* is that it is pointed to by the biggest hubs in the network, like the broadcasters *Antena3.com* and *Atresplayer.com*, amongst others. These results allow us to answer our first research question.

As mentioned, we have analysed the extent to which the observed authority scores are different from what we would expect if the hyperlinks of the news network were sent randomly—significant scores are labelled * in Figure 1. Interestingly, though, we have also assessed the significance of the difference in means of authority scores of new and legacy news outlets. According to our results using bootstrapping techniques for non-independent observations, the type of news provider is not significant to explain levels of authority in the online domain (see Figure A3 in Supplementary File). In other words, the individual significant differences previously identified do not persist at the category level.

Additionally, in Figure 1 we can also see that the reach of the news outlets, i.e., the total amount of visits media receive, is only slightly associated with authority scores (rho = 0.30, p-value < 0.05). Hence, the number of visits that one outlet receives does not offer a complete explanation for its media authority score. Rather, the re-

sults suggest it is the underlying structure of connections that confer legacy media with authority. Their ability to generate valuable content attracts other media's attention and grants them the most central positions in the news network.

By means of the hyperlink-induced topic search algorithm, we have also identified hubs in the news network, i.e., pointers to useful information. This analysis reveals that there are more digital-born media acting as pointers to useful news sources, i.e., hubs, than as authorities. Among them, we find *Libertadigital.com*, *Elconfidencial.com*, and *Diarioinformacion.com*. Some news outlets, mainly broadcasters like *Antena3.com*, *Atresplayer.com* and *Lasexta.com*, play a double role in the news domain; not only are they relevant authorities but they also specialize in identifying the most important news sources on the web.

To answer our second research question—to find out which type of news provider is in control of the audience flow—we have measured the weighted betweenness centrality of the outlets included in our study. More precisely, we have determined the differences in the brokerage power of legacy and digital-born media across demographics. Figure 2 shows the results of the analysis. As suggested in this analysis, the control of the flow of audiences is almost a monopoly of the legacy organizations. They hold top brokerage positions in each age group.

Yet, when we drill down to young audiences' navigation patterns, those aged 18 to 24, the analysis reveals that digital-born media obtain, on average, a higher betweenness centrality score than in other
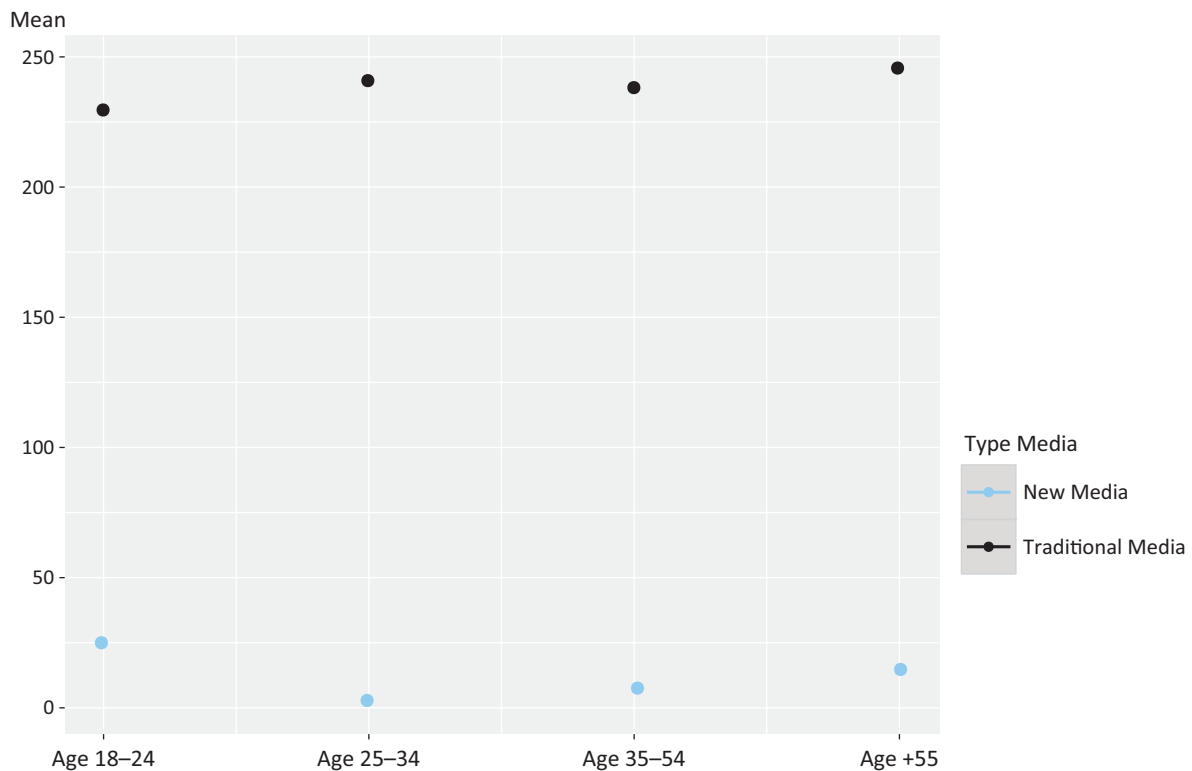

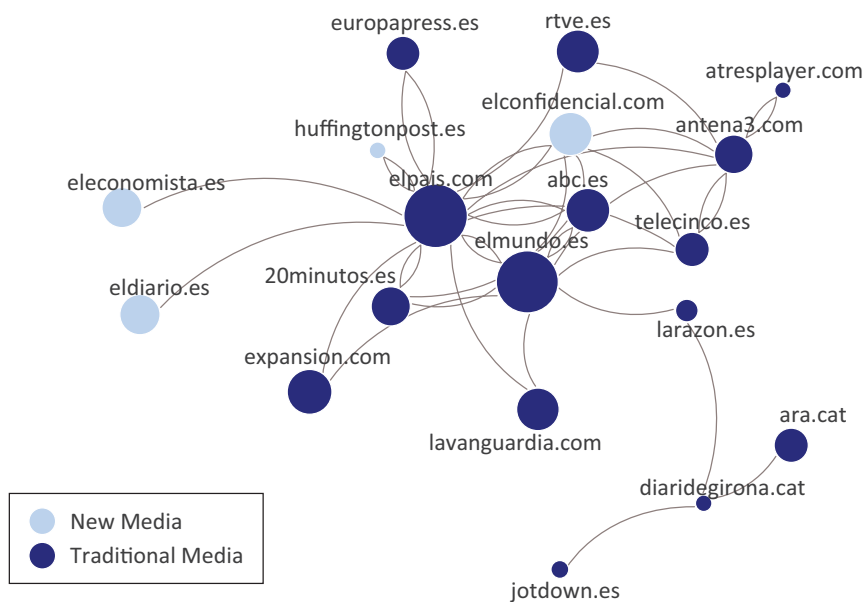**Figure 2.** Mean of betweenness centrality scores in the audience networks by different age groups.

Figure 3. Audience brokers in the young audience network.

age groups. According to our results, three digital-born outlets are playing the role of top brokers for young audiences: *Elconfidencial.com*, *Eldiario.es* and *Eleconomista.es*. Figure 3 maps their positions within the strongest connected component, i.e., the central part of the young audience network. The figure shows the most important brokers. Here, nodes represent news media outlets and their size is proportional to their brokerage power. To obtain the strongest connected component of the structure, we have iteratively removed the weakest connections of the young audience network, a process known as percolation in network science (Borge-Holthoefer & Gonzalez-Bailon, 2015). The advan-

tage of this method is that it reveals the areas with the highest levels of audience overlapping and helps us to better understand patterns of news navigation between new and legacy media.

Overall, and in line with previous research, we can see that patterns of young audience navigation differ from those of the general public (see Figure 4). More specifically, and to answer our final research question, young people confer to digital-born sites a central role in their media diets. These types of news sources compete for the control of young audience flow with legacy brands. They occupy core positions within the network and have major roles in the mediation of news attention.
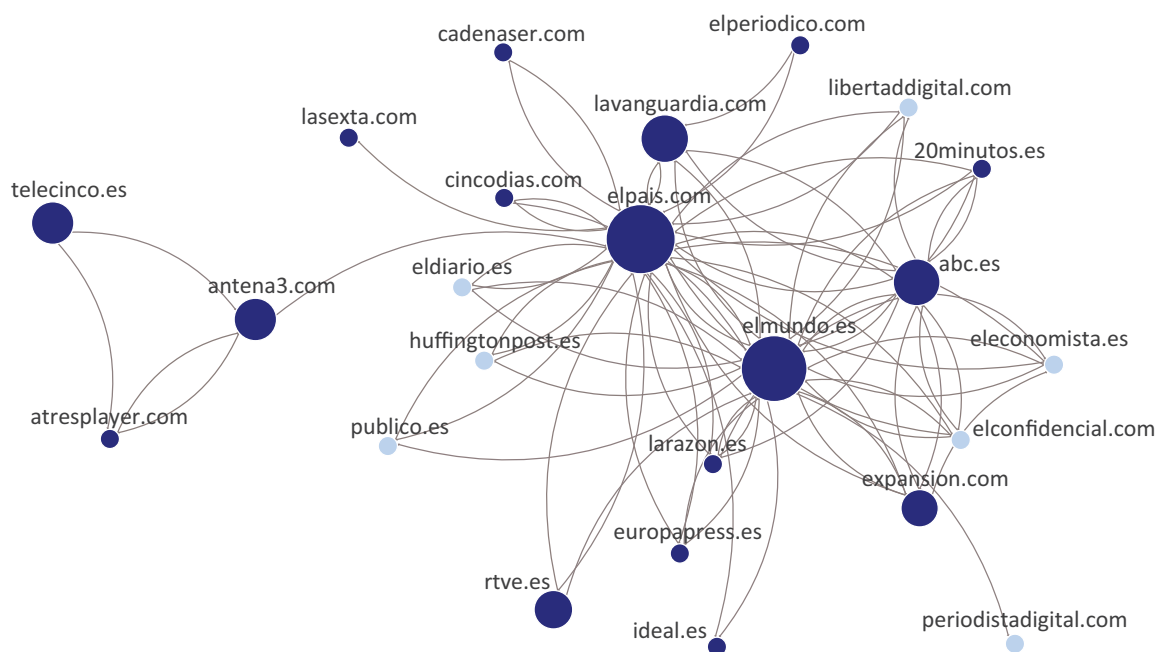
Figure 4. Audience brokers in the general audience network.

Of note, and to properly interpret these results, one must take into consideration that the audience networks represent aggregated news consumption behaviour. Hence, they do not provide the actual path that people follow when navigating more than two consecutive sites in the news domain. Nonetheless, it does provide the best representation of the likeliest path that at the aggregated level news users will follow.

Finally, Figure 4 shows the distribution of brokerage power in the general audience network, and it is congruent with our previous findings. It shows the greater role of legacy outlets—see the greater size of dark-blue coloured nodes—as mediators of audience behaviour in the online news domain. In light of this evidence, we argue that offline heritage is still a key element to understanding media centrality in the news domain.

5. Conclusion

This study highlights that a negotiation of roles among digital-born outlets and legacy media is taking place in the online news domain. We provide evidence that legacy media are regarded as the most authoritative sources and that their role does not hinge on their reach. Also, legacy media retain much of the control of the audience flow. However, young audiences, who exhibit distinctive news consumption patterns, confer brokerage power to native digital outlets. Hence, these new entrants occupy central positions for the control of the audience flow too.

Previously, the literature had already highlighted tensions between new digital actors and legacy news organizations (Carlson, 2007, 2017; Singer, 2003; Tandoc & Jenkins, 2017). Before the advent of the web, it noted and assessed the efforts in journalistic practice to delimitate the boundaries of the profession and to assert its authority as qualified sources for news information (Zelizer, 1990). Narratives techniques were used by journalists to enhance their authority and preserve the centrality of legacy media in the information flow. Undoubtedly, the control of the means of production and distribution favoured that goal. As noted elsewhere, the profession wanted to limit the access of new actors and preserve the monopoly of the authority of legacy media (Singer, 2003). Interestingly, this same rhetorical technique that was used by legacy media later served digital-born outlets to assert their authority and the uniqueness of the news content they started offering (Carlson & Usher, 2016; Park, 2009).

However, the web soon lowered the cost of the production and distribution of news. Legacy media then faced a hard reality: limiting and controlling the entrance of new players was not a realistic goal. Mainly because, worldwide, audiences have recognized the value of digital-born outlets which have been keener on embracing innovation and differentiating themselves from legacy players not only through narratives but also modes of participation. Interestingly, audiences

have legitimated new processes to produce and distribute information brought by digital-born outlets and, thereby, have conferred authority to digital-born outlets (Carlson, 2017).

The extent to which these changes have challenged the monopoly of legacy organizations, though, was not previously tackled in our field from an empirical perspective. Here we have proposed an observational approach to fill this gap. In this study, building on a multidimensional approach for measuring media authority, we have argued that the role of most prominent news media in the online domain hinges on the dominance of authoritative and broker positions in the news and audience network respectively. We have mapped the network structure between news providers, which has offered a useful lens to interpret the actual underlying relationships among news outlets based on their hyperlink structure. Moreover, we have assessed the extent to which audiences have legitimized the new entrants and accepted their journalistic knowledge (Carlson, 2017)—and whether, in doing so, the monopoly of the centrality of the legacy media has been challenged.

Overall, the contribution of our study is twofold. First, it offers an empirical framework that can be applied regardless of the media context to unveil influential relationships in the media ecology. In doing so, it helps to advance our understanding of how the Internet impacts the role of legacy media. Second, this article contributes to the literature on the reconfiguration of the media ecosystem by providing direct evidence on the role of legacy media as compared to that of digital-born outlets. Our results suggest that the latter are still far from displacing traditional brand outlets, whose stronger offline reputation may be securing their central role in the online news domain. Yet the youngest audiences are exerting their influence as catalysts for change.

It has been beyond the scope of this study to compare at the macro-level the structure of the online news domain. Yet previous research, focused on the overall web, suggests that regional and linguistic differences might be shaping the production and consumption of news (Taneja, 2016; Wu & Ackland, 2014). In fact, a visual inspection of the news audience network would suggest it too. Future research can build upon this work to provide an evidence-based answer to these questions.

Finally, some limitations deserve consideration. Our data represents a snapshot of the news domain, and a temporal analysis might be necessary to account for variations in the positions of the news media across time. Additionally, future studies exploring the structural mechanisms underlying betweenness centrality should also consider more refined measures of audience attention than total audience visits, e.g., total minutes spent on news content. Regardless of these, we believe this study makes a significant empirical contribution to the study of the media's changing roles in the online news domain.

COGITATIO

Conflict of Interests

The authors declare no conflict of interests.

Supplementary File

Supplementary material for this article is available online in the format provided by the authors (unedited).

References

Althaus, S. L., & Tewksbury, D. (2000). Patterns of Internet and traditional news media use in a networked community. *Political Communication*, *17*(1), 21–45.

Anderson, C., Bell, E., & Shirky, C. (2012). *Post-industrial journalism: Adapting to the present*. New York, NY: Columbia Journalism School.

Asociación para la Investigación de Medios de Comunicación. (2014). *Estudio general de medios: Resumen abril 2013 a marzo 2014* [General study of media: Executive summary April 2013 to March 2014]. Madrid: Asociación para la Investigación de Medios de Comunicación. Retrieved from http://reporting.aimc.es/index.html#/main/cockpit

Bale, P., Walmsley, D., Wustermann, L., Ericson, M., & Barber, G. (2018). Indicator: Citations and references. *The Trust Project*. Retrieved from https://docs.google.com/document/d/10QQuXvle8bT3hqRiGCaato6qhO_UwaM3-bsUdTsO2j0/edit#bookmark=id.w3k8dc4i7mzi

Bennett, W. L. (2003). New media power. In N. Couldry & J. Curran (Eds.), *Contesting media power* (pp. 17–37). Lanham, MD: Rowman & Littlefield.

Borge-Holthoefer, J., & Gonzalez-Bailon, S. (2015). Scale, time, and activity patterns: Advanced methods for the analysis of online networks. In N. Fielding, R. Lee, & G. Blank (Eds.), *Handbook of online research methods* (pp. 259–276). Thousand Oaks, CA: Sage Publications.

Borodin, A., Roberts, G. O., Rosenthal, J. S., & Tsaparas, P. (2005). Link analysis ranking: Algorithms, theory, and experiments. *ACM Transactions on Internet Technology (TOIT)*, *5*(1), 231–297.

Brogi, E., Ginsborg, L., Ostling, A., Parcu, P. L., & Simunjak, M. (2015). *Monitoring media pluralism in europe: Testing and implementation of the media pluralism monitor 2015*. San Domenico di Fiesole: European University Institute.

Cagé, J. (2016). *Saving the media: Capitalism, crowdfunding, and democracy*. Cambridge, MA: Harvard University Press.

Cage, J., Viaud, M.-L., & Herve, N. (2015). The production of information in an online world. *The Review of Economic Studies*. https://doi.org/10.1093/restud/rdz061

Caño, A. (2016, March 3). Carta abierta del director de El País a la redacción del periódico [Open letter from el país director to the newsroom]. *El Pais*. Retrieved from http://elpais.com/elpais/2016/03/03/actualidad/1457031570_991358.html

Carlson, M. (2007). Blogs and journalistic authority: The role of blogs in US election day 2004 coverage. *Journalism Studies*, *8*(2), 264–279.

Carlson, M. (2017). *Journalistic authority: Legitimating news in the digital era*. New York, NY: Columbia University Press.

Carlson, M., & Usher, N. (2016). News startups as agents of innovation: For-profit digital news startup manifestos as metajournalistic discourse. *Digital Journalism*, *4*(5), 563–581.

Castells, M. (2007). Communication, power and counter-power in the network society. *International Journal of Communication*, *1*(1), 238–266.

Castells, M. (2009). *Communication power*. New York, NY: Oxford University Press.

Chadwick, A. (2013). *The hybrid media system: Politics and power*. Oxford: Oxford University Press.

Chung, C. J., Nam, Y., & Stefanone, M. A. (2012). Exploring online news credibility: The relative influence of traditional and technological factors. *Journal of Computer-Mediated Communication*, *17*(2), 171–186.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, *94*, 95–120.

Coleman, R., & McCombs, M. (2007). The young and agenda-less? Exploring age-related differences in agenda setting on the youngest generation, baby boomers, and the civic generation. *Journalism & Mass Communication Quarterly*, *84*(3), 495–508.

Colino, A. (2013). Evolución de la financiación de los medios de comunicación en España: ¿Hacia un modelo más sostenible? [Financial evolution of the spanish media: Towards a more sustainable model?] *Papeles de Europa*, *26*(1), 46–69.

Cook, T. E. (1998). *Governing with the news: The news media as a political institution*. Chicago, IL: University of Chicago Press.

Cooper, R. (1996). The status and future of audience duplication research: An assessment of ratings-based theories of audience behavior. *Journal of Broadcasting & Electronic Media*, *40*(1), 96–111.

Couldry, N., & Curran, J. (2003). *Contesting media power: Alternative media in a networked world*. Lanham, MD: Rowman & Littlefield Publishers.

Croteau, D. (2006). The growth of self-produced media

content and the challenge to media studies. *Critical Studies in Media Communication*, *23*(4), 340–344.

Downie, L., & Schudson, M. (2009). The reconstruction of American journalism. *Columbia Journalism Review*. Retrieved from https://archives.cjr.org/reconstruction/the_reconstruction_of_american.php

Dutton, W. H., Gillett, S. E., McKnight, L. W., & Peltu, M. (2004). Bridging broadband Internet divides: Reconfiguring access to enhance communicative power. *Journal of Information Technology*, *19*(1), 28–38.

Fenton, N. (2010). *New media, old news: Journalism and democracy in the digital age*. New York, NY: Sage Publications.

Fortunati, L., Sarrica, M., O'Sullivan, J., Balcytiene, A., Harro-Loit, H., Macgregor, P., . . . De Luca, F. (2009). The influence of the Internet on European journalism. *Journal of Computer-Mediated Communication*, *14*(4), 928–963.

Franklin, M. [heyitsfranklin2]. (2018, June 29). Hi friends, I'm interrupting your regularly scheduled program to say: I have some personal news. After almost 5 years, today is my last day at @Mashable. Starting next week, I'm joining the team at @nytimes as a social strategy editor! [Tweet]. Retrieved from https://twitter.com/heyitsfranklin2/status/101281592984206 1313?s=19

Freeman, L. C. (1979). Centrality in social networks conceptual clarification. *Social Networks*, *1*(3), 215–239.

Garnham, N. (1995). The media and the public sphere. *Intermedia*, *14*(1), 28–33.

Gonzalez-Bailon, S., & Wang, N. (2013). The bridges and brokers of global campaigns in the context of social media. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2268165

Goodhardt, G. J., & Ehrenberg, A. S. C. (1969). Duplication of television viewing between and within channels. *Journal of Marketing Research*, *6*(2), 169–178.

Goodhardt, G. J., Ehrenberg, A. S. C., & Collins, M. A. (1987). *The television audience: Patterns of viewing* (2nd ed.). Gower House: Gower Publishing Co. Ltd.

Guess, A. M. (2015). Measure for measure: An experimental test of online political media exposure. *Political Analysis*, *23*(1), 59–75.

Gurevitch, M., Coleman, S., & Blumler, J. G. (2009). Political communication: Old and new media relationships. *The ANNALS of the American Academy of Political and Social Science*, *625*(1), 164–181.

Hallin, D. C., & Mancini, P. (2004). *Comparing media systems: Three models of media and politics*. Cambridge: Cambridge University Press.

Harlow, S., & Salaverría, R. (2016). Regenerating journalism: Exploring the "alternativeness" and "digitalness" of online-native media in Latin America. *Digital Journalism*, *4*(8), 1001–1019.

Hermida, A., & Thurman, N. (2008). A clash of cultures: The integration of user-generated content within professional journalistic frameworks at British newspaper websites. *Journalism Practice*, *2*(3), 343–356.

Interactive Adversiting Bureau. (2011, October 27). *Comscore, seleccionado para la medición de la audiencia online en España* [ComScore is apointed as the official audience meter in Spain] [Press Release]. Retrieved from https://www.comscore.com/esl/prensa-y-eventos/comunicados-de-prensa/2011/11/comscore-selected-as-online-media-measurement-currency-in-spain?cs_edgescape_cc=us

Jarvis, J. (2016, February 14). Stop the presses! The independent is dead: Long live the independent. *Buzzmachine*. Retrieved from https://buzzmachine.com/2016/02/14/stop-presses

Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)*, *46*(5), 604–632.

Kleinberg, J. M., Kumar, R., Raghavan, P., Rajagopalan, S., & Tomkins, A. S. (1999). The web as a graph: Measurements, models, and methods. *Computing and Combinatorics*. https://doi.org/10.1007/3-540-48686-0_1

Kosslyn, J., & Yu, C. (2017). Fact Check now available in Google Search and News around the world. *Google*. Retrieved from https://blog.google/products/search/fact-check-now-available-google-search-and-news-around-world

Ksiazek, T. B. (2011). A network analytic approach to understanding cross-platform audience behavior. *Journal of Media Economics*, *24*(4), 237–251. https://doi.org/10.1080/08997764.2011.626985

Larsson, A. O. (2013). Staying in or going out? Assessing the linking practices of Swedish online newspapers. *Journalism Practice*, *7*(6), 738–754.

Lee, J. K. (2007). The effect of the Internet on homogeneity of the media agenda: A test of the fragmentation thesis. *Journalism & Mass Communication Quarterly*, *84*(4), 745–760.

Lewis, S. C. (2012). From journalism to information: The transformation of the Knight Foundation and news innovation. *Mass Communication and Society*, *15*(3), 309–334.

Livingstone, S. (1999). New media, new audiences? *New Media and Society*, *1*(1), 59–66.

Majó-Vázquez, S., Zhao, J., & Nielsen, R. K. (2017). The digital-born and legacy media news media on Twitter during the French presidential elections. Oxford: Reuters Institute. Retrieved from https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2017-07/Maj%C3%B3-V%C3%A1zquez%20-%20The%20Digital-Born%20and%20Legacy%20News%20Media%20on%20Twitter.pdf

Majó-Vázquez, S., Zhao, J., Nurse, J. R., & Nielsen, R. K. (2017). Digital-born and legacy news media on Twitter during the UK general election. Oxford: Reuters Institute. Retrieved from https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2017-11/Digital-Born%20and%20Legacy%20News%20Media%20UK%20Factsheet.pdf

Mansell, R. (2004). Political economy, power and new media. *New Media and Society*, *6*(1), 74–83.

McChesney, R. W., & Pickard, V. (2011). *Will the last reporter please turn out the lights: The collapse of journalism and what can be done to fix it*. New York, NY: The New Press.

Meraz, S. (2009). Is there an elite hold? Traditional media to social media agenda setting influence in blog networks. *Journal of Computer-Mediated Communication*, *14*(3), 682–707. https://doi.org/10.1111/j.1083-6101.2009.01458.x

Meraz, S., & Papacharissi, Z. (2013). Networked gatekeeping and networked framing on# Egypt. *The International Journal of Press/Politics*, *18*(2), 138–166.

Meyer, P. (2004). The influence model and newspaper business. *Newspaper Research Journal*, *25*(1), 66–83.

Minder, R. (2015, November 5). Spain's news media are squeezed by government and debt. *The New York Times*. Retrieved from http://www.nytimes.com/2015/11/06/world/europe/as-spains-media-industry-changes-rapidly-some-worry-about-objectivity.html?_r=0

Mitarbeiter fordern Revolution von unten [Employees demand evolution from below]. (2016). *Southwest Broadcasting*. Retrieved from http://www.swr.de/swr2/kultur-info/krise-beim-spiegel-mitarbeiter-fordern-revolution-von-unten/-/id=9597116/did=17196744/nid=9597116/yexa7/index.html

Mitchell, A., Gottfried, J., Shearer, E., & Lu, K. (2017). *How Americans encounter, recall and act upon digital news*. Washington, DC: Pew Research Center. Retrieved from http://assets.pewresearch.org/wp-content/uploads/sites/13/2017/02/08183209/PJ_2017.02.09_Experiential_FINAL.pdf

Napoli, P. (2008). Hyperlinking and the forces of massification. In J. Turow & L. Tsui (Eds.), *The hyperlinked society: Questioning connection in digital age* (pp. 56–69). Ann Arbor, MI: The University of Michigan Press.

Newman, M. (2010). *Networks: An introduction*. Oxford: Oxford University Press.

Newman, N., Fletcher, R., Kalogeropoulos, A., & Nielsen, R. K. (2019). *Digital news report 2019*. Oxford: Reuters Institute for the Study of Journalism.

Nicholls, T., Shabbir, N., & Nielsen, R. K. (2016). *Digital-born news media in Europe*. Oxford: Reuters Institute for the Study of Journalism.

Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, *32*(3), 245–251.

Park, D. W. (2009). Blogging with authority: Strategic positioning in political blogs. *International Journal of Communication*, *3*, 250–273.

Park, S., Ko, M., Lee, J., & Song, J. (2013). *Agenda diversity in social media discourse: A study of the 2012 Korean general election*. Paper presented at the 7th International Conference on Web and Social Media, Boston, MA, USA.

Pavlik, J. V. (2001). *Journalism and new media*. New York, NY: Columbia University Press.

Pew Research Center. (2015). *State of the news media 2015*. Washington, DC: Pew Research Center.

Press Association of Madrid. (2013). *Informe anual de la profesion periodistica 2013* [Annual report of journalistic profession 2013]. Madrid: Press Association of Madrid.

Press Association of Madrid. (2015). *Informe anual de la profesion periodística* [Annual report of journalistic profession]. Madrid: Press Association of Madrid.

Prior, M. (2009). The immensely inflated news audience: Assessing bias in self-reported news exposure. *Public Opinion Quarterly*, *73*(1), 130–143.

Riordan, K. (2014). *Accuracy, independence, and impartiality: How legacy media and digital natives approach standards in the digital age*. Oxford: Reuters Institute for the Study of Journalism. Retrieved from https://reutersinstitute.politics.ox.ac.uk/publication/accuracy-independence-and-impartiality

Robinson, S. (2007). "Someone's gotta be in control here": The institutionalization of online news and the creation of a shared journalistic authority. *Journalism Practice*, *1*(3), 305–321.

Sagarra, O., Font-Clos, F., Pérez-Vicente, C. J., & Díaz-Guilera, A. (2014). The configuration multi-edge model: Assessing the effect of fixing node strengths on weighted network magnitudes. *EPL (Europhysics Letters)*. https://doi.org/10.1209/0295-5075/107/38002

Salaverría, R. (2016). *Ciberperiodismo en Iberoamérica: Marco general* [Cyber journalism in Ibero-America: A general framework]. Madrid: Fundación Telefónica and Editorial Ariel.

Salaverría, R., Sádaba, C., Breiner, J. G., & Warner, J. C. (2019). A brave new digital journalism in Latin America. In M. Túñez-López, V.-A. Martínez-Fernández, X. López-García, X. Rúas-Araújo, & F. Campos-Freire (Eds.), *Communication: Innovation and quality* (pp. 229–247). Berlin: Springer.

Schoepp, S. (2016, March 23). Reporteros en el salón [Journalist ins the livingroom]. *Ctxt.Es.* Retrieved from http://ctxt.es/es/20160323/politica/4979/reporteros-salon-periodismo-españa-comunicación-viva-la-prensa-escrita-documentos-ctxt.htm

Singer, J. B. (2003). Who are these guys? The online challenge to the notion of journalistic professionalism. *Journalism*, *4*(2), 139–163.

Singer, J. B. (2007). Contested autonomy: Professional and popular claims on journalistic norms. *Journalism Studies*, *8*(1), 79–95.

Smith, B. (2020, March 1). Why the success of The New York Times may be bad news for journalism. *The New York Times*. Retrieved from https://www.nytimes.com/2020/03/01/business/media/ben-smith-journalism-news-publishers-local.html?smid=nytcore-ios-share

Suárez, E. (2020). *How to build a successful subscription news business: Lessons from Britain and Spain*. Oxford: Reuters Institute for the Study of Journalism.

Retrieved from https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2020-02/Eduardo_Suarez_Digital_subscriptions_fellowship_paper.pdf

Tandoc, E. C., Jr., & Jenkins, J. (2017). The Buzzfeedication of journalism? How traditional news organizations are talking about a new entrant to the journalistic field will surprise you! *Journalism*, *18*(4), 482–500.

Taneja, H. (2016). Mapping an audience-centric World Wide Web: A departure from hyperlink analysis. *New Media & Society*, *19*(9), 1331–1348.

Taneja, H., & Webster, J. G. (2016). How do global audiences take shape? The role of institutions and culture in patterns of web use. *Journal of Communication*, *66*(1), 161–182.

Taneja, H., Wu, A. X., & Edgerly, S. (2018). Rethinking the generational gap in online news use: An infrastructural perspective. *New Media & Society*, *20*(5), 1792–1812.

The Independent becomes the first national newspaper to embrace a global, digital-only future. (2016, February 12). *The Independent*. Retrieved from http://www.independent.co.uk/news/media/press/the-independent-becomes-the-first-national-newspaper-to-embrace-a-global-digital-only-future-a6869736.html

Top sites in Spain. (2015). *Alexa*. Retrieved from http://www.alexa.com/topsites/countries/es

Weber, M. S., & Monge, P. (2011). The flow of digital news in a network of sources, authorities, and hubs. *Journal of Communication*, *61*(6), 1062–1081.

Webster, J. G., & Ksiazek, T. B. (2012). The dynamics of audience fragmentation: Public attention in an age of digital media. *Journal of Communication*, *62*(1), 39–56. https://doi.org/10.1111/j.1460-2466.2011.01616.x

Wills, A. (2014). The full New York Times innovation report. *Scribd*. Retrieved from http://www.scribd.com/doc/224608514/the-full-new-york-times-innovation-report

Wu, L., & Ackland, R. (2014). How Web 1.0 fails: The mismatch between hyperlinks and clickstreams. *Social Network Analysis and Mining*, *4*(1), 1–7.

Zelizer, B. (1990). Achieving journalistic authority through narrative. *Critical Studies in Media Communication*, *7*(4), 366–376.

About the Authors

Sílvia Majó-Vázquez is a Research Fellow at the Reuters Institute for the Study of Journalism of the University of Oxford. Her research areas include news audience behaviour, digital news structure and the role of new and legacy media in the online news domain. She was Research Fellow at DiMeNet group at Annenberg School for Communication at University of Pennsylvania in 2015 and 2016. She was Visiting Scholar at the Data Science Center of Universitat Pompeu Fabra in 2019.

Ana S. Cardenal is Professor of Political Science at the Open University of Catalonia and affiliated to the Universitat Pompeu Fabra. She was a Visiting Scholar at the Center for Latin American Studies at Stanford University and a Fulbright Fellow at the Department of Politics at New York University. Her current research interests include digital audiences, digital media and political attitudes, public opinion formation, effects of the political information environment on parties and voting, and computational (and research) methods.
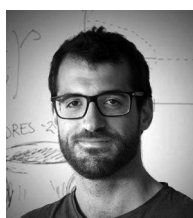
Oleguer Sagarra has a PhD in Complex Systems and Human Mobility by University of Barcelona and MSc in Computational Physics by Politechnical University of Catalonia. After a research stay at Massachusetts Institute of Technology (MIT) he founded Dribia, a data science innovation studio based in Barcelona, where he acts as co-CEO and Lead Data Scientist. His research interests include human mobility, synchronization, complex systems, citizen science and complex networks.

Pol Colomer de Simón has a PhD in Structure of Complex Networks by University of Barcelona and MSc in Computational Physics by Politechnical University of Catalonia. After a research stay in Center for Applied Internet Data Analysis at San Diego (CAIDA) he founded Dribia, a data science innovation studio based in Barcelona, where he acts as co-CEO and Lead Data Scientist. His research interests include topology of complex networks, epidemiology and phase transitions.