International Journal of Communication 11(2017), 5068–5092 1932-8036/20170005

Imitation in the Quest to Survive: Lessons from News Media on the Early Web

MATTHEW S. WEBER
KATHERINE OGNYANOVA
Rutgers University, USA

ALLIE KOSTERICH
Pace University, USA

This article examines the patterns of hyperlinking among key online newspapers from 1996 to 2000 and provides critical insight into the processes by which media companies adapt to new technology. Theories of organizational imprinting and imitation in the media industry are used to frame the rise of online news in an effort to describe processes of growth and to track the interactions among legacy newspapers during a formational period in the development of online news. Patterns of digital connectivity reveal the evolution of an increasingly close-knit online news community and the trajectory of leadership positions in the online environment. The analysis reveals various approaches that leading organizations used as they adapted to online technology, providing guidance for organizations moving forward.

Keywords: news, organizations, new media, Internet, network analysis

The newspaper industry is an archetype of the struggle to adapt to new technology, as newspapers have faced serious challenges in transforming to digital production (Boczkowski, 2004). Although the World Wide Web was invented in 1990, the revolution in mass communication did not commence until 1994, when a beta version of Netscape enabled graphic Web browsing. In turn, this article focuses on 1996 to 2000 as a critical period of adaptation for newspapers on the Web. In examining this process of adaptation, scholars have studied the degree of interactivity afforded by online newspapers (Massey & Levy, 1999), the growth of online news business models (Thompson & Wassmuth, 2001), and the efficiency of online channels for communicating news in a timely manner (Li, 2006). Studies also looked at the evolution of content on the Web, including an examination of the rapid growth of online news content (Greer & Mensing, 2004) and a subsequent analysis of the changes in technology and services offered by those newspapers (Greer & Mensing, 2006).

Matthew S. Weber: matthew.weber@rutgers.edu
Katherine Ognyanova: katya.ognyanova@rutgers.edu
Allie Kosterich: akosterich@pace.edu
Date submitted: 2016–07–14

In contrast, this work explores theoretical foundations of adaptation to change to explain how online newspapers adapt to new technologies. We draw on *imprinting*, a construct borrowed from organizational science, to explain the decisions newspapers made as they moved online. To evaluate the early dynamics of growth on the Web, we examine patterns of hyperlinking among the websites of legacy newspapers in the 1990s. Legacy newspapers, especially those on the early Web, are important because of their central role in the industry. The core question we address is how legacy newspapers developed a digital presence through the use of hyperlinks as structural tools for establishing and maintaining positions in the online ecosystem. Moreover, through the frame of imprinting, we address how online newspapers learned from one another during this formative period, as evidenced by imitation of structural patterns of organizations founded during a common time period.

The following sections interweave history of the Web with the rise of online news and the decline of legacy newspapers. Echoing other histories of the Web (e.g., Ankerson, 2012), we use an analysis of archived Internet data to capture the transition of legacy newspapers to the Web. The analysis provides a macro-level perspective of industry-wide change using historical data sources to contextualize the process. We conclude by discussing lessons that can be applied to future generations of technology innovations.

The Newspaper Industry and the Emergence of the Web

The U.S. newspaper industry thrived for more than a century as a stable population of organizations and encountered several technological disruptions, including changes in paper manufacturing that slashed printing costs and increased production (Lee, 1937) and the advent of the Linotype, which changed printing standards (Schwarzlose, 1987). Technological changes chipped away at the solid foundation of the newspaper industry, but newspaper circulation grew through the early 1990s. Revenue and circulation began to slip in the mid-1990s and dropped dramatically in the 2000s.

Technological changes, however, were not the sole factor influencing transformation in the news industry at this time (Deuze, 2007). Market conditions were accompanied by other social forces that influenced newspapers' online development. Boczkowski (2004), for example, highlights the variations in organizational structure and practices as an important factor in newsroom adoption of technology. Gilbert (2005) adds that the threat of competition among these organizations influenced technological investment processes, highlighting the importance of social forces in examining the process of technological adoption.

In the move to adapt, newspaper content was made available through bulletin-board systems and online service providers such as Prodigy and America Online as early as 1992. By 1994, approximately 60 U.S. newspapers offered content via an Internet access point; however, fewer than 10 were accessible on the Web (Li, 2006). The Raleigh *News & Observer* launched a Web page in September 1994, followed by the *San Francisco Examiner/Chronicle* in November, and the *San Jose Mercury News* in December. *The New York Times* launched its Web edition in early 1996 (Gunter, 2003).

According to Li (2006), 1995 and 1996 marked a turning point for online newspapers. Prior to this point, the goal of many newspapers was simply to claim that they had a presence online. Beginning in

1996, newspapers began to experiment with the development of unique content and the use of the Web for distributing timely information. In 1996, about 15% of newspapers had some form of online presence—amounting to approximately 250 online newspapers. As of 1997, half of U.S. newspapers (approximately 745 news outlets) had websites, and by 1998, there were 1,749 online newspapers in the U.S. (Editor & Publisher Directories, 2006). By 2000, online news was developing as an integral factor in the business model of newspaper companies.

The Role of the Hyperlink

The hyperlink is a basic unit of communication on the Web (Halavais, 2008) and an element that has an important impact on digital journalism. It has existed since the Web was first established and was part of the Hypertext Markup Language (HTML) specifications for the first Web pages (Berners-Lee, 1999). Hyperlinks helped to create the basic structure of a website, guiding users between Web pages and across domains. Most early online newspapers contained lists of article headlines hyperlinked to Web pages containing full articles (Scott, 2005). From the start, hyperlinks were an integral part of the Web's structure, central to information flow across Web pages. The degree of hyperlinking used to guide readers, and to point to other content sources, proved to be a critical factor in driving the growth of newspapers (Weber, 2012). The use of hyperlinks by U.S. online newspapers increased considerably between 1997 and 2001 (Tremayne, 2004). Patterns of hyperlink use over time provide insight into the means by which newspapers integrated new Web technology into their workflows. The degree of hyperlink use also provides insights into the structure of and interactions within the online newspaper community.

In a review of extant hyperlink scholarship, De Maeyer (2012) argues that existing research aligns with one of two axes of thought: The first axis frames hyperlinks as an indicator of social phenomena, whereas the second axis frames hyperlinks as an indicator of structural patterns. These two related bodies of research demonstrate that hyperlinks are both social and structural. As a social phenomenon, hyperlinks serve as a means for exchanging resources and information (Gonzales-Bailon, 2009), expressing individual interests and preferences (Etling, Kelly, Faris, & Palfrey, 2009), and conveying political conversation (Adamic & Adar, 2001). In the context of journalism, hyperlinks are indicators of gatekeeping (Dimitrova, Connolly-Ahern, Williams, Kaid, & Reid, 2003) and credibility (Tremayne, 2005). C. W. Anderson (2013) argued that hyperlinks are a critical source for understanding change in journalism, as the changing nature of their use is indicative of broader systemic changes in the newspaper industry.

From a structural perspective, hyperlinks serve as proxies for endorsements (Kleinberg, 1999) and alliances (Park, Kim, & Barnett, 2004), and they can be signals of shared characteristics or markers of reputation (Ognyanova & Monge, 2013). Hyperlinks provide key infrastructure for the navigational process by which users move through a website and experience its content (Brügger, 2009). At a global level, scholars have found that broad hyperlinking patterns reveal common structures across topical domains, such as the existence of highly interconnected clusters of Web pages that provide cores for communities of Web pages (Barabasi, 2002). Hyperlinks also provide a means for mapping the global distribution of information (Tremayne, 2004).

Shumate and Lipp (2008) argue that a better understanding of structural patterns over time is critical to understanding the foundational building blocks of organizational action. In the context of this article specifically, De Maeyer and Le Cam (2015) argue for a deeper examination of journalism's hyperlinking history as a key to interpreting the present role of hyperlinks in the news industry and to understanding ongoing change processes. In turn, this article focuses primarily on the structure of hyperlinking activity during early periods of newspaper Web use, as opposed to examining the context and content surrounding given hyperlinks. Nevertheless, because hyperlinks guide consumers from the content of one Web page to that of another, or from one location on a Web page to another, these structural patterns are broadly indicative of the structure of the content itself.

Hyperlinking Patterns of Online Newspapers

In response to the need for continued research examining the structural role of hyperlinks in the development of news media, this work explores patterns of imitation and imprinting through the structural relationships represented by hyperlinks. Hyperlinks proved central to the development of online newspapers but created a challenge for news producers as hyperlinks allowed for free-flowing access to information. In contrast with editorial policies, much of the early history of the Web was marked by democratic collaboration, such as the sharing of information and knowledge among website developers. In the early days of online news, the use of hyperlinking to drive traffic to other sources was viewed as anathema (Boczkowski, 2004). Editors saw hyperlinking as a problem because it encouraged the reader to leave the newspaper's own page.

The hyperlinking patterns of a website are often directly indicative of the type of information provider (Weber & Monge, 2011). As De Maeyer and Holton (2016) found, hyperlinks between news websites have stronger connections than those in other contexts, as journalists place greater significance on the transmission of information between sources. Early patterns of hyperlinking on the Web represent a significant experiment with new technology by an industry that was traditionally reluctant to adapt to new technological regimes (Coddington, 2014). Furthermore, early decisions regarding online presence can have an impact on the growth trajectory of a website (Weber, 2012). The relationship of websites, as demonstrated by the hyperlink structure of their connections, similarly impacted the evolution and success of online newspapers over time. Thus, the structure of relationships between newspapers on the early Web is interrogated as follows:

RQ1: What is the pattern of hyperlink connections between online newspapers in the early years of their emergence on the Web?

Imprinting and Structural Imitation

The early structure of online newspapers was largely dependent on the role of Web technology as perceived by innovators in the 1990s. The organizational theory of imprinting provides a basis for exploring early perceptions of the Web and for understanding the impact of those perceptions on the growth of online news. Imprinting posits that during critical founding and transition periods, organizations are particularly susceptible to external stimuli, such as the influence of other organizations (Marquis &

Tilcsik, 2013). Imprinting is the degree to which a new organization, or an organization undergoing transformation, reflects the economic, political, and social conditions during its founding or transition period (Johnson, 2007). Imprinting provides an explanation for the high degree of similarity among organizations in an industry: Organizations develop similar characteristics in response to the shared pressures and demands to which they are subjected. The imprinting framework is applied in a wide variety of fields (for a review, see Marquis & Tilcsik, 2013) and is well suited for the study of adaptation under new conditions.

Imprinting as a Process

In recent years, research on imprinting and imitation during the foundational periods of organizations focused on the process by which imprinting is realized (Baron, Burton, & Hannan, 1996; Johnson, 2007). Johnson suggests this process can be broken into its component elements, focusing first on the process of founding and the way in which "available elements" are incorporated into new structures, and second, on understanding how these elements are then reproduced over time. Work in this area has demonstrated that the environmental conditions during founding have a strong impact on the trajectory of an organization and its ability to remain competitive (Shinkle & Kriauciunas, 2012).

Imprinting and Imitation in Journalism

Within journalism, imprinting may be linked to a trend toward imitation in both form and content (Boczkowski, 2010). Davis and Greve (1997) note that imitating is a key practice during periods of change and has been well established as a process for organizational learning (Huber, 1991). Indeed, the act of imitation has been viewed in several studies as part of the early process of imprinting (Boeker, 1987; Kim, Yang, & Hwang, 2002; Kogut, 1993).

Within the context of journalism, the advent of the Web, uncertainty in the media industry, and the accelerated news cycle made the practice of reproducing existing materials increasingly central (Boczkowski, 2009). Patterns of imitation were also evident in a study of online news production in Korea, which found that the 13 most popular online newspapers delivered similar headlines (Lim, 2012). Even design similarities are not "random happenstance," but emerge from a variety of forces in a dynamic environment (Cooke, 2005). This echoes prior work on intermedia agenda setting showing that newspapers would look to industry leaders and would often duplicate what was being covered on competitors' news digests (Boyle, 2001; Meraz, 2011).

Although the analysis of imitation in news typically examines content, this article focuses on similar tendencies in the hyperlinking patterns of online newspapers as another indicator of imitation. This echoes De Maeyer's (2012) view that hyperlinks are both social and structural in nature and loosely draws on imprinting literature that focuses on how founding conditions lead to similarities in structural patterns (Marquis, 2003; Milanov & Fernhaber, 2009). Imitation in the use of structural hyperlinking patterns may thus be an indicator of a broader process of imprinting resulting from the environmental conditions at founding and from interactions with a larger cohort of organizations.

Organizations, such as traditional newspapers, entering into a new space tend to engage in a process of organizational bricolage. Through bricolage, organizations borrow ideas and practices from competitors and adapt based on previously successful strategies (Perkmann & Spicer, 2014). This is akin to the manner in which newspapers experimented with new forms of online publishing. The more organizations rely on a common set of resources, the more likely those organizations are to adopt similar practices (Chen & O'Mahony, 2009). Resources, in this context, include knowledge of a new technology, human capital, and technological standards. The broad focus on hyperlinking patterns thus looks at the development of common practices and access to common resources.

Reproducing and repurposing news stories are understood practices, but less is known about the similarities in structural features among online newspapers. Similarities in structural features are likely to be reflected by the degree to which Web adaptation strategies of legacy newspapers mirror each other in terms of hyperlinking. As discussed, imprinting is prominent in periods of uncertainty and leads to similarities across organizations. In this case, the focus is on similarities in hyperlinking behavior, as follows:

RQ2: Do online newspapers share similar hyperlinking patterns in the early years of the Web?

Method

We use hyperlink analysis to examine the structure of relationships among newspapers. Hyperlink analysis has been established as a key approach to studying the growth of websites (Park & Thelwall, 2003) and is a critical tool for understanding news and information flows (Weber & Monge, 2011; Weber & Nguyen, 2015). Social network analysis is used as a methodological approach to examine the structure of the newspaper industry on the Web.

Data

This article leverages data from the Internet Archive (https://archive.org), a nonprofit organization that aims to archive and preserve the Web. We used data from the Internet Archive's 20th Century collection, which is the largest archive of Web pages from 1996 to 2000. No coherent archive of Web pages prior to 1996 is available. To facilitate analysis, archived data were extracted in the Web Archive Transformation (WAT) metadata format. WAT is a metadata specification that captures key fields from raw Web pages and allows for faster analysis and data extraction. Key fields include inbound hyperlinks, outbound hyperlinks, file size, keywords, descriptive text, and content type (HTML, images, video, etc.). The compressed records for the 20th Century collection are 7.2 TB. Data were sorted and extracted into tabular format using a customized data extraction tool and a high-performance computing cluster (Weber & Nguyen, 2015). Data and source code used in this study are available online.[1] Each record contains the originating Web page (source), the linked-to page (destination), the date of the record, the number of hyperlinks between the two pages on that date, and any associated descriptive text.

[1] Data and source code used in this study are available at https://github.com/mwe400/IJOCNewsOnEarlyWeb

Data Selection

This research focused on mainstream news sites that existed during the early period of the Web. Mainstream news websites—mainly a few key competitors— produced the majority of news content during this period (Weber & Monge, 2011). Our analysis was restricted to the online behavior of the top 25 newspapers in terms of print circulation in each year from 1996 through 2000. Annual circulation data were collected from the Alliance for Audited Media. Table 1 shows the cumulative list of the 28 newspapers occupying positions in the top 25 at any time during that time period. Print circulation was used as opposed to online circulation for two primary reasons: First, print circulation is a key indicator of revenue during this period, and second, online measures of audience were limited during the early development of the Web. Online start date was estimated through a triangulation of sources, including first recorded date in the Internet Archive. The structural relationships of the focal online newspapers were extracted from a data set of more than 1.2 billion Web pages, which is a repository covering a much larger data set of newspaper content. The data were reduced to a smaller scale to represent key relationships between major actors.

Hyperlink Data

The extracted data included all hyperlinks to and from the domains of the 28 legacy newspapers studied (N = 3,524,364 records). The key properties of each hyperlink were its source, destination, weight, and timestamp. The source and destination were specific domains (e.g., "nyt.com," "yahoo.com"). The weight was the total number of hyperlinks from any page in the source domain to any page in the destination domain.

Because of the frequency of Web crawls used to archive the data, a yearly time frame was used for analysis. Summary data of the crawling activity is presented in Table 2. As the table shows, there was a significant variance in the amount of crawling activity. *The Atlanta Journal-Constitution* was crawled once in 1996, whereas *The New York Times* was crawled 257 days in 2000. Website domains were crawled an average of four days in 1996, and an average of 174 days in 2000. The low frequency of crawling activity in 1996 is due to the nascent nature of the Web at that time; the crawling activity is far more significant by 2000. Our focus on crawling activity across a given year helps to account for variance within the year, and by averaging the data, we are able to overcome the impact of differences in crawl frequencies.

Table 1. Summary Statistics for Focal Newspapers.

Newspaper	Founding	Online Founding	Avg. Circulation	Avg. of Outlinks	Avg. of Webites Outlinked	Avg. of Inlinks	Avg. of Webites Inlinked
The Atlanta Journal-Constitution	2001[a]	1998	341,481	30	18	1,666	1,253
The Arizona Republic	1890	1995	455,457	785	527	3,810	2,439
The Boston Globe	1872	1998[b]	471,279	11	2	153	124
Chicago Sun Times	1948	1997	488,171	471	299	8,405	5,323
Chicago Tribune	1847	1998	655,843	800	400	3,882	2,556
The Dallas Morning News	1885	1996	509,008	156	63	4,413	2,922
The Denver Post	1892	1996	365,295	1,096	877	4,322	2,814
Detroit Free Press	1831	1996	369,346	352	196	6,063	4,210
Houston Chronicle	1901	1996	548,767	1,560	1,228	8,468	5,331
Los Angeles Times	1881	1996	1,088,989	1,454	836	18,546	12,188
Miami Herald	1903	1998	353,274	29	17	598	426
Minneapolis Star Tribune	1867	1996	371,010	597	352	5,635	3,724
New York Daily News	1919	1996	730,627	144	16	1,820	1,156
New York Post	1801	1998	432,214	525	29	1,816	999
Newsday	1940	1996	568,881	794	596	6,079	3,366
The New York Times	1851	1996	1,294,596	194	120	40,055	25,009
The Orange County Register	1905	1999	362,350	700	607	2,375	1,600
The Oregonian	1850	1997	351,577	989	679	2,891	1,695

The Philadelphia Inquirer	1829	1999	418,064	224	124	477	335
The Plain Dealer	1842	1996	392,107	265	114	2,338	1,464
Rocky Mountain News	1859	1998	402,767	48	28	702	248
The San Diego Union Tribune	1868	2002	376,341	—	—	—	—
San Francisco Chronicle	1865	1996	467,597	1,075	721	21,206	11,523
The Star Ledger	1832	1996	405,130	475	218	10,317	6,323
Tampa Bay Times	1884	1998	358,014	116	70	185	63
USA Today	1982	n/a	1,693,185	3,226	1,737	61,584	33,109
The Wall Street Journal	1889	1996	1,809,191	667	507	16,519	12,424
The Washington Post	1877	n/a	807,586	4,502	2,544	31,862	19,529

[a] *The Atlanta Journal* and *The Atlanta Constitution* formally merged into one daily newspaper, *The Atlanta Journal-Constitution*, in November 2001. The two newspapers had shared a newsroom since 1982. The merger date is given here as the founding, although the two newspapers ran a joint website as of 1998. *The Atlanta Constitution* was founded in 1868, and *The Atlanta Journal* was founded in 1883.
[b] December 12, 1998.

Table 2. Crawling Activity Summary.

	Days Crawled				
Newspaper	1996	1997	1998	1999	2000
The Atlanta Journal-Constitution	2	1	—	2	13
The Arizona Republic	3	14	10	107	219
The Boston Globe	—	—	1	6	10
Chicago Sun Times	—	5	1	63	132
Chicago Tribune	—	—	—	113	210
The Dallas Morning News	—	4	10	108	215
The Denver Post	—	1	7	78	184
Detroit Free Press	1	3	2	99	197
Houston Chronicle	1	3	20	137	236
Los Angeles Times	5	10	—	127	252
Miami Herald	—	—	—	7	15
Minneapolis Star Tribune	1	10	14	106	219
New York Daily News	—	—	3	11	62
New York Post	—	—	1	78	249
Newsday	1	6	10	96	200
The New York Times	4	5	10	82	216
The Orange County Register	—	—	—	87	178
The Oregonian	—	1	15	139	254
The Philadelphia Inquirer	—	—	—	36	181
The Plain Dealer	1	14	20	106	241
Rocky Mountain News	—	—	5	22	19
The San Diego Union Tribune	1	6	28	72	157
San Francisco Chronicle	4	3	6	140	229
The Star Ledger	1	37	34	145	257
Tampa Bay Times	—	—	1	21	50
USA Today	13	19	20	138	230
The Wall Street Journal	8	26	25	112	226
The Washington Post	10	36	20	62	206

When a hyperlink between two domains was recorded more than once during a given year (62% of all cases), those hyperlinks were combined and their weight was averaged. The average was used as opposed to a sum to capture the relative strength of the relationship during the time period, as some domains were crawled and recorded more often than others. The five-year data set (n = 869,128 records) contained 211,741 unique domains, including the focal newspaper websites along with every site that had hyperlinks from or to them. Table 3 provides descriptive statistics for the hyperlinks among the focal websites.

Table 3. Hyperlink Descriptives for the 28 Newspapers Included in the Analysis, 1999–2000.

	1996	1997	1998	1999	2000
Total number of unique hyperlinks	22,141	107,116	157,168	227,960	354,743
Total number of unique websites	12,623	45,867	68,155	95,221	142,750
Unique hyperlinks to newspaper domains	14,112	94,427	152,089	217,093	330,881
Unique hyperlinks from newspaper domains	8,071	12,757	5,133	10,990	24,053
Unique hyperlinks among newspaper domains	42	68	54	123	191
Unique hyperlinks among newspaper domains (w > 1)	18	49	35	76	98
Unique websites linking to newspaper domains	6,730	39,117	65,081	89,295	129,601
Unique websites linked from newspaper domains	6,605	9,536	4,492	8,916	18,864

Analysis

Analyses were conducted using the R platform for statistical computing (R Core Team, 2015; RStudio, 2015). To address the first research question on emerging hyperlink patterns among newspapers, we extracted the annual networks of hyperlinks across news sites. To reduce potential noise in the data, we dropped ties with an average weight of one (a single hyperlink between any pages in the domains recorded in the respective year). We computed a range of network characteristics for that graph for each time point including density, reciprocity, transitivity, and in-degree and out-degree centralization.

Density refers to the number of ties that exist as a ratio of the total potential number of ties. *Reciprocity* is a tendency toward a back-and-forth exchange, or symmetric social interaction. Reciprocity is measured as the proportion of hyperlinks from Site A to Site B where the reverse hyperlink from B to A is also present. *Transitivity*, or clustering, is based on the propensity to close triangles and is linked to the emergence of closely knit social groups. Transitivity is computed as the proportion of cases where existing hyperlinks from A to B and A to C are also accompanied by a hyperlink from B to C. Finally, *centralization* measures the extent to which certain key sites in the hyperlink network tend to have a disproportionately large number of incoming hyperlinks (in-degree centralization) or outgoing hyperlinks (out-degree centralization). High in-degree centralization signals the emergence of dominant actors that accumulate influence and have more connections compared with the rest of the network (Wasserman & Faust, 1994).

To evaluate the observed levels of reciprocity, transitivity, and centralization, we conducted conditional uniform graph (CUG) tests with 1,000 replications each, conditioning on network size and density (B. S. Anderson, Butts, & Carey, 1999). In those tests, networks are drawn randomly from a uniform distribution of graphs with the same number of nodes and edges as the observed data. The characteristics of the generated graphs are compared with those of the empirical network. This allows researchers to assign more meaningful interpretations to features of the observed data. For example, we might not know if a centralization score of 0.30 is low, high, or typical for a certain kind of network. A CUG test might show that out of thousands of generated graphs with the same size and density as ours, very few or none have a centralization score as high as 0.30, pointing instead to the existence of some underlying network formation mechanisms that favor centralization.

To address the second research question on similarities in linking patterns across newspapers, we examined incoming and outgoing hyperlink profiles of each news domain. For each year in the data, we identified the top 5,000 domains with the most hyperlinks from online newspapers and the top 5,000 with the most hyperlinks to online newspapers. We refer to the vector describing the number of hyperlinks each of those 5,000 sent to a news domain as that domain's *incoming hyperlink profile*. The *outgoing hyperlink profile* of an online newspaper was the vector describing the volume of its outgoing hyperlinks to each of the top 5,000 websites. We selected the top 5,000 websites in each case so that the results are less likely to be influenced by the growing total number of sites from year to year and are more likely to be determined by the specific portfolio of websites that were linked most frequently.

To examine hyperlinking similarity, we calculated the correlations between outgoing hyperlink profiles of all online newspapers in the data. To evaluate similarity, we computed the correlations between the incoming hyperlink profile vectors for each online newspaper. We further examined the average profile correlations over time. To evaluate the observed average similarity levels, we used permutation tests similar to the CUG test. We reshuffled the incoming and outgoing hyperlinks of online newspapers 1,000 times and compared the average profile similarity scores from the resulting permuted profiles to those of the observed data.

Results

The descriptive analysis of the hyperlink data provides a snapshot of the general patterns of newspaper growth on the Web. Table 4 summarizes the growth of hyperlinking for online newspapers in our sample.

Table 4. Summary of Hyperlinking Behavior by Year.

Year	Newspapers with External Links	Outgoing Hyperlinks (*n* Hyperlinks)		Incoming Hyperlinks (*n* Hyperlinks)		Outgoing Hyperlinks (*n* Websites)		Incoming Hyperlinks (*n* Websites)	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
1996	14	833	1,704	1,321	2,106	475	958	614	881
1997	18	1,055	1,734	7,714	12,962	580	951	3,632	5,379
1998	23	382	631	10,150	16,711	205	354	5,633	8,474
1999	27	657	786	12,536	17,981	407	605	8,040	11,155
2000	27	1,250	1,557	17,909	22,898	859	1,081	12,255	15,571

In general, out-linking and in-linking activity grew over time, as did the number of unique websites that engaged in hyperlinking activity. This is expected given the general growth of the Web over time, and in particular during the early period of the Web. There is a notable decline in 1998. The decrease in 1998 was due, in part, to changes in Internet standards, which interfered with existing archiving protocols. The result was that the archival records for 1998 did not increase in accordance with the growth of the Web. Figure 1 provides a summary visualization of the interaction among the 28 focal websites and illustrates the increased hyperlinking activity among them over time and the increase in clustering as domains became more connected.

The hyperlink network of the focal online newspapers was sparse but grew denser over time. The network exhibited very little reciprocity in the early years of the Web. In the conditional uniform graph test, the proportion of graphs with reciprocity as high or higher than the observed data was 100% for 1996 and 1998, 15% for 1997, 83% for 1999, and 3% for 2000. These results show that it was not until 2000 that the hyperlink structure started exhibiting more reciprocity than expected from a network of that size and density.

The transitivity of the online newspaper domain network also increased over time. The zero transitivity in 1996 was due to the small number of ties in the network; this is not surprising given that hyperlinks were first used for internal structure of websites. From the 1,000 graphs with a similar density generated in our test for that year, 90% also had a transitivity of zero. From 1997 to 2000, none of the networks generated for the CUG test had transitivity as high as the observed measure. After its tentative first year on the Web, the newspaper network began exhibiting considerable triadic closure (clustering).

The network became more centralized from 1996 to 1999, both in in-degree and out-degree centralization. In 2000, the centralization scores dropped compared with 1999 but remained relatively high. None of the 1,000 graphs drawn for each year in the data from a density-conditioned uniform graph distribution had in-degree or out-degree centralization as high as that of the empirical networks. As centralization increased over time, key actors emerged within the newspaper hyperlink network. *The New York Times* and *The Washington Post* received the most incoming hyperlinks from other publications in the later years under examination. The most active website in terms of linking to other news sources was the *Detroit Free Press*.

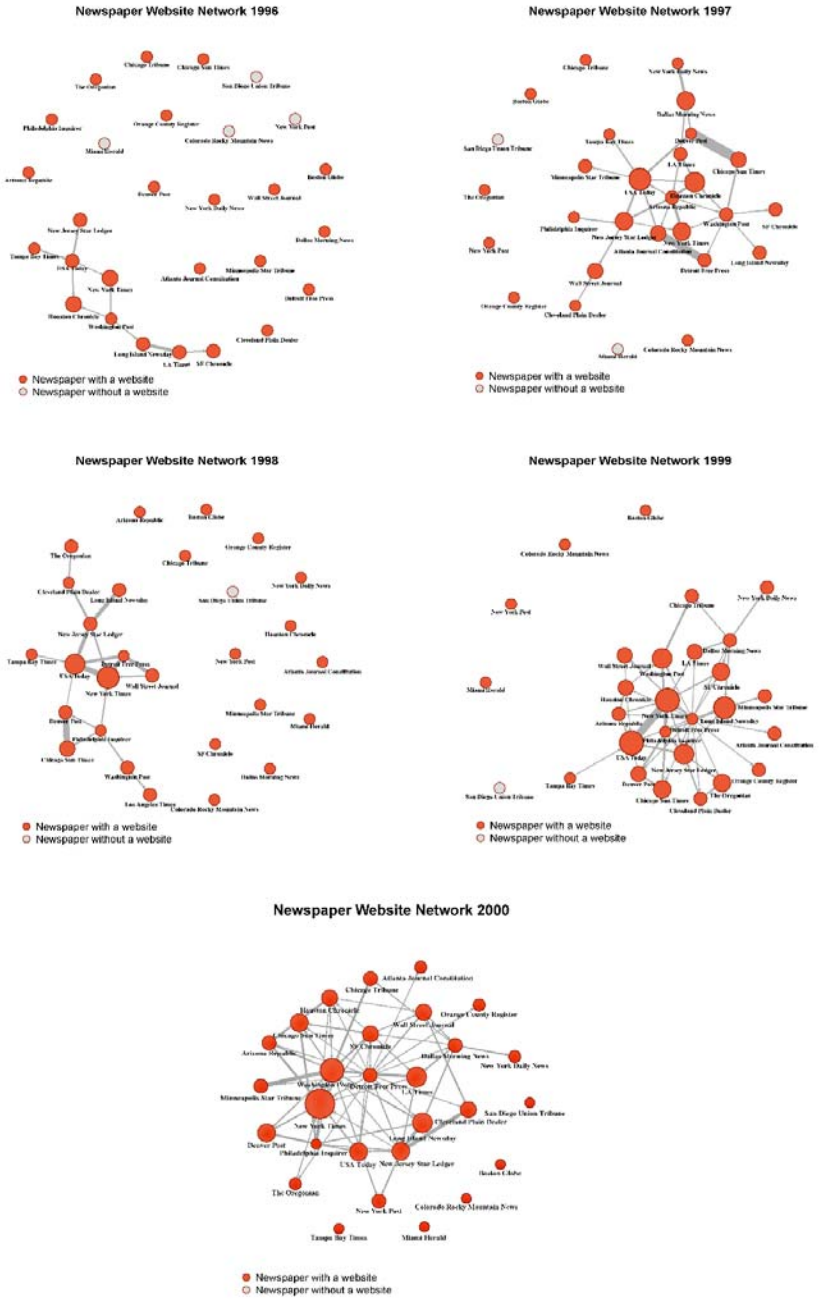

Figure 1. Hyperlink networks among news organizations (hyperlinks with weight larger than 1).

To address the second research question, we examined the similarity in the incoming and outgoing hyperlinks of the news domains with the top 5,000 most active sites in our sample each year. Our analysis focused on top-level domains; in other words, we examined inbound and outbound links from nytimes.com, as opposed to internal links within the domain. Figure 2 depicts the average outgoing hyperlink profile correlations, which ranged between 0.08 and 0.12 over the five years in the data. We also examined the similarity of the online newspapers based on the profiles of incoming hyperlinks for each online newspaper. Correlations ranged from 0.04 to 0.09, and were highest in the beginning and the end of the examined period.

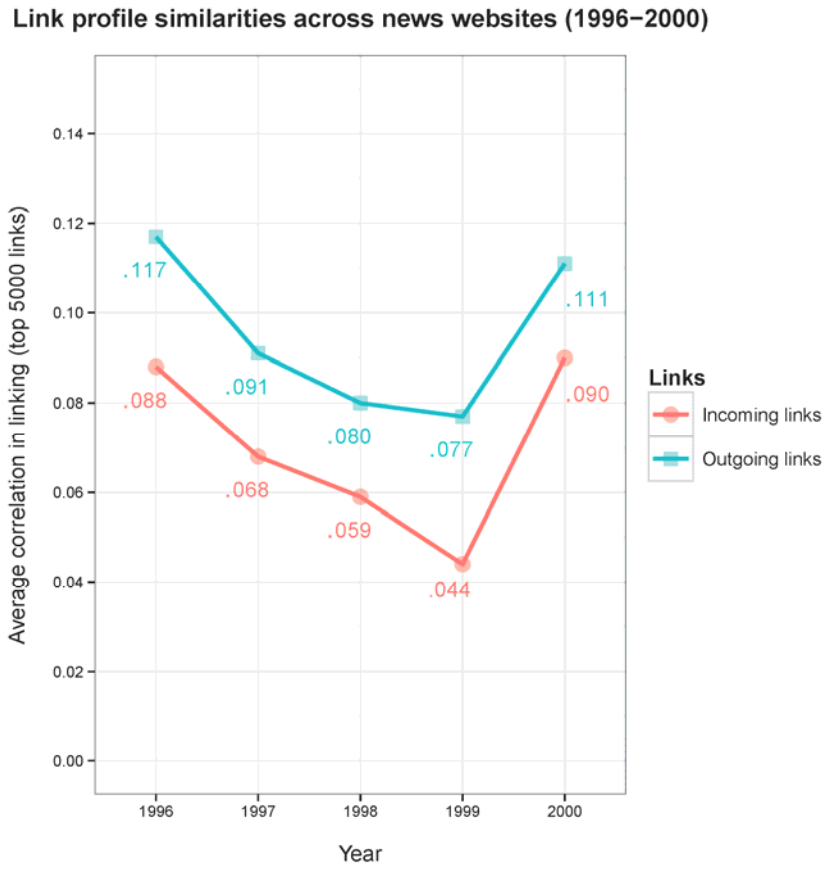

Figure 2. Average hyperlink profile similarity among news websites by year.

Because of the fairly large and diverse number of domains with hyperlinks to and from news organizations, the hyperlink profile correlations appear fairly small in size. Permutation tests show, however, that given the number of websites and the density of the ties, the observed similarities are highly significant. Reshuffling the incoming and outgoing hyperlinks of online newspapers 1,000 times for each time point, we found that in all but one case, 100% of the permutations for both incoming and outgoing hyperlinks produced average profile similarities lower than those of the observed data. The exception was outgoing hyperlinks in 1996, where 98.5% of the permutations had average correlations lower than the observed data.

Discussion

This study details patterns by which early online newspapers emerged on the Web and points to implications about imitation among newspaper companies. The general upward trend in hyperlinking patterns reflects an increasing diversity of news content, the growing traffic of users necessitating more sophisticated navigation to different sources of content, and the maturing technological competencies within the industry. Quantitatively, these findings reinforce certain expectations that have been put forth in prior scholarship. For instance, interpreted as a high-level structural indicator (Barabasi, 2002), the hyperlink patterns revealed here trace the development of an increasingly interconnected online newspaper community. As a marker of reputation, the increase in the degree of hyperlinking to specific Web domains further reveals the emergence of leaders in a new digital space. For example, as illustrated in Table 2, *USA Today* received a high degree on inbound hyperlinking activity, suggesting many other websites included links pointing to the domain of this newspaper. *USA Today* was an early adopter of hyperlinks, and recent work has shown that newspapers that adopted hyperlinking practices early on were more likely to emerge as leaders in the online news space (Weber & Monge, 2011).

Finally, from an evolutionary perspective (Shumate, 2012), structural similarities in linking provide suggestive clues about imprinting and imitation in news organizations. Here, patterns of outbound hyperlinking are most critical, as they are indicative of the actions of a specific newspaper. Inbound hyperlinking activity gives a broader context characterizing the behavior of websites within this ecosystem, but the outbound hyperlinks are most representative of the actions of stakeholders within a given newspaper.

Descriptive findings of link volume detail the overall activity within the industry during its online transition. Beginning in 1996, newspapers were shifting to digital production, but in general they were relatively inactive and, as illustrated by the low density of their hyperlinks, isolated on the Web. Over time, density of the network increased, pointing to a higher degree of exchange and interaction among websites in the news community. Reciprocity remained low, showing that hyperlinking among the websites in this analysis was generally unidirectional. Although newspapers occasionally linked to one another early on, hyperlinks between websites were seldom reciprocated. This aligns with knowledge about early uses of hyperlinking for advertising and information purposes and is consistent with the patterns discussed below, signaling the emergence of online leaders.

Hyperlinks as Indicators of Structure and Content

Another important perspective is the use of hyperlinks as status cues—an approach used early on by Google's search-result-ranking algorithm that allowed it to outperform all of its competitors. As demonstrated in Table 1, there was a wide variation in hyperlinking activity across websites. The consistently high centralization of the newspaper network suggests that even among that elite group of publications, some, such as *The New York Times* and *The Washington Post*, were particularly prominent online. Others, such as the *Detroit Free Press*, were especially likely to generate hyperlinks to other media. These differences map onto two distinct online strategies driving the structural patterns of early newspaper websites. Whereas some publications were focused on maximizing user retention by providing more content and having fewer opportunities for their readers to exit from a domain via outgoing hyperlinks, others saw themselves as hubs directing people to a variety of available information sources. This pattern of hubs and authorities was still nascent on the early Web, but recent research demonstrates its increasingly popularity over time (Weber & Monge, 2011).

For major newspapers, the use of hyperlinks steadily increased over time. This broad trend is demonstrated in Table 3, where the increase in outbound hyperlinking activity is evident. Year over year, there was a significant growth in the connections among newspaper domains as the average number of incoming and outgoing hyperlinks of each news outlet grew. Interpreting this increase in connectivity through the lens of newsroom norms and practices is consistent with prior research (Gao & Vaughn, 2006) suggesting that hyperlinking was increasingly recognized as a strategic tool. Hyperlinks provide structure to a website but also direct attention to particular content by guiding the consumer to a given Web page (Carpenter, 2010; Ketterer, 2001).

Exploring larger structural patterns enables an examination of the trajectory of online newspapers as a dynamic system of organizational actors. As legacy newspapers adapted to the Web, they began to accept and realize the importance of interconnectedness among websites. Again, this pattern of connectivity is evident in the results presented in Table 3, which illustrates the increase in activity outbound and also between newspaper domains included in this study. Connectivity helps to increase information flow, but it also serves institutional purposes (e.g., signaling affiliation or joint action). For example, newspapers under common owners formed content-sharing alliances and were likely to link to one another under such agreements. The news-site community became more closely knit over time, as evidenced by the growing transitive closure of the hyperlink network. Another relevant finding showing changing perceptions in the industry is the fact that reciprocity increased in 2000, marking a shift in the hyperlinking strategies of newspaper companies on the Web.

Hyperlinking During a Foundational Period

Hyperlinking is a key aspect of digital adaptation. The investigation of structural similarities among newspaper websites supports the notion that imprinting occurred in the industry during its early online years. In this context, imprinting is studied by examining the correspondence in hyperlinking patterns across newspapers. Following studies on entrepreneurs during early foundational periods, this takes a structural view of imprinting and emphasizes the process of structural emergence during this key

period of change (Milanov & Fernhaber, 2009). While our tests did not focus on establishing causal mechanisms, one plausible explanation for similarity in this context is the trend toward imitation across news organizations.

Previous research has demonstrated the major impact of the Web in promoting newsroom imitation and content replication (Boczkowski, 2010) and more broadly in newsrooms competing for consumer attention (McManus, 1994). There is evidence that isomorphism has become an entrenched mechanism of growth for online news worldwide (Lim, 2012). In particular, it is reasonable to expect that normative influences of the isomorphism process (Sydow, Schreyögg, & Koch, 2009) led newspaper companies to adopt similar hyperlinking strategies. Certain practices of hyperlinking were established by the leading newspapers and quickly adopted by competitors in an attempt to keep up. The similarity in these patterns is shown in the correlation results provided in the online supplemental material.[2]

Interestingly, similar patterns are at play in today's online news environment. In 2015, Facebook launched its Instant Articles feature, whereby content of partner newspapers is prominently featured on the social media company's application and loads faster than the content of competitors. The benefits of this tight coupling of content partners with Facebook have prompted an increase in partnerships with the social media company (Gaines, 2016).

The role of technology in promoting content imitation is well established (Boczkowski, 2010; Mecca & Mumford, 2014), and this study contributes to the literature by establishing the dynamic as a broader structural phenomenon that clearly influenced the development of news media on the early Web. Our analysis examined similarities in the outgoing hyperlinks generated by news websites (indicating strategic connectivity to the rest of the Web) and the incoming hyperlinks (signaling how presence was perceived by others online). More recent work has further indicated that social media technology may propagate trends of imitation by allowing competitors to more closely monitor one another's activity (Chadha & Wells, 2016).

Both earlier research focusing on content and the present study examining hyperlink structure suggest that imprinting and imitation occur during periods of disruption and as new organizational structures are emerging. Imprinting may occur at the cost of innovation and competitiveness, but it may also help an organization survive periods of disruption (Marquis & Tilcsik, 2013). Consistent with imprinting, the degree of hyperlinking similarity declined over time from 1996 to 1999, only to increase again in 2000. The decrease is partly explained by the general increase in the number of websites with hyperlinks to and from news organizations as years passed. It may also be attributable to increasing comfort with Web technology.

These findings also highlight the importance of imprinting for innovation within the newspaper industry, echoing research from entrepreneurship literature (Hsu & Lim, 2013). As legacy newspapers enter mobile and social media, they face opportunities and challenges not unlike those encountered in the initial transition to the Web. Similar to practices familiar from early online newspapers, those endeavors

[2] Online supplemental material is available at https://github.com/mwe400/IJOCNewsOnEarlyWeb

initially focused on broadcasting and popularizing journalistic content. Use is evolving beyond the simple sharing of news stories as mobile apps and social media become ubiquitous among mainstream news sources. News media outlets are once again under economic pressure to fully exploit the available new platforms and formats. Technological and market forces have led many outlets toward increased interactivity with other information sources and a more personalized and responsive approach to audiences.

Limitations

The present study has three key limitations. First, the number of longitudinal variables available for inclusion in this study was relatively limited. In general, there is a lack of complete data detailing the organizational attributes of news organizations in prior decades. For example, many of these hyperlinks could be determined by other patterns of influence such as ownership and advertising partnerships; however, a cohesive record of this type of data is scarce.

Second, the use of archival Internet data is a limitation. Archivists have documented numerous challenges with archiving digital content. For instance, an ongoing Library of Congress initiative to capture election websites has faced challenges associated with capturing content across technological standards (Grotke, 2011). Prior studies examining the completeness of hyperlink data from the Internet Archive demonstrated that as much as 20% to 30% of hyperlinks are missing from any given subset (Weber & Nguyen, 2015). Given that such omissions tend not to be systematic, it is arguable that the resulting data represent a random sampling of hyperlinking behavior.

Third, the period examined in this study covers five years. By comparison, many studies of imprinting focus on longer time periods. Admittedly, a longer time period would allow for a better understanding of the degree to which early imitation carried on into later cohorts, thus focusing on the second stage of the process of imprinting suggested by Johnson (2007). Notably, other studies of emergence and entrepreneurship (Baron et al., 1996) have used shorter time periods. Some recent studies have used time periods between five and 10 years (Baron et al., 1996; Heiman & Clarysse, 2005; Leung, Foo, & Chaturvedi, 2013).

Future Studies

This study underscores how imprinting factored into the early formation of online newspapers. At present, new technology is once again creating new avenues of competition among Web services. The rise of new access points to the Internet, such as mobile applications and social media websites, is driving a new process of reinvention for media companies. Findings from this study suggest that there will be substantial levels of imitation and similarity early in the adaptation of new technology. Future research should examine the impact that levels of structural, technological, and content imitation have on the performance and success of organizations over time.

References

Adamic, L. A., & Adar, E. (2001). *You are what you link*. Paper presented at the 10th annual International World Wide Web Conference, Hong Kong.

Anderson, B. S., Butts, C. T., & Carey, K. (1999). The interaction of size and density with graph-level indices. *Social Networks*, *21*(3), 239–267. doi:10.1016/S0378-8733(99)00011-8

Anderson, C. W. (2013). What aggregators do: Towards a networked concept of journalistic expertise in the digital age. *Journalism*, *14*(8), 1008–1023. doi:10.1177/1464884913492460

Ankerson, M. S. (2012). Writing Web histories with an eye on the analog past. *New Media & Society*, *143*(3), 384–400. doi:10.1177/1461444811414834

Barabasi, A.-L. (2002). *Linked: The new science of networks*. Cambridge, MA: Perseus.

Baron, J. N., Burton, M. D., & Hannan, M. T. (1996). The road taken: Origins and evolution of employment systems in emerging companies. *Industrial and Corporate Change*, *5*(2), 239–275. doi:10.1093/icc/5.2.239

Berners-Lee, T. (1999). *Weaving the Web*. New York, NY: HarperOne.

Boczkowski, P. J. (2004). *Digitizing the news: Innovation in online newspapers*. Cambridge, MA: MIT Press.

Boczkowski, P. J. (2009). Technology, monitoring, and imitation in contemporary news work. *Communication, Culture, & Critique*, *2*, 39–59. doi:10.1111/j.1753-9137.2008.01028.x

Boczkowski, P. J. (2010). *News at work: Imitation in an age of information abundance*. Chicago, IL: University of Chicago Press.

Boeker, W. (1987). Strategic origins: Entrepreneurial and environmental imprinting at founding. *Academy of Management Proceedings*, *1987*(1), 150–154. doi:10.5465/AMBPP.1987.17534046

Boyle, T. P. (2001). Intermedia agenda setting in the 1996 presidential election. *Journalism & Mass Communication Quarterly*, *78*(1), 26–44. doi:10.1177/107769900107800103

Brügger, N. (2009). Website history and the website as an object of study. *New Media & Society*, *11*(1–2), 115–132. doi:10.1177/1461444808099574

Carpenter, S. (2010). A study of content diversity in online citizen journalism and online newspaper articles. *New Media & Society*, *12*(7), 1064–1084. doi:10.1177/1461444809348772

Chadha, K., & Wells, R. (2016). Journalistic responses to technological innovation in newsrooms: An exploratory study of Twitter use. *Digital Journalism*, *4*(8), 1020–1035. doi:10.1080/21670811.2015.1123100

Chen, K. K., & O'Mahony, S. (2009). Differentiating organizational boundaries. In B. G. King, T. Felin, & D. A. Whetten (Eds.), *Studying differences between organizations: Comparative approaches to organizational research* (Vol. 26, pp. 183–200). Bingley, UK: Emerald Group.

Coddington, M. (2014). Normalizing the hyperlink. *Digital Journalism*, *2*(2), 140–155. doi:10.1080/21670811.2013.785813

Cooke, L. (2005). A visual convergence of print, television, and the Internet: Charting 40 years of design change in news presentation. *New Media & Society*, *7*(1), 22–46. doi:10.1177/1461444805049141

Davis, G. F., & Greve, H. R. (1997). Corporate elite networks and governance changes in the 1980s. *American Journal of Sociology*, *103*(1), 1–37. doi:10.1086/231170

De Maeyer, J. (2012). Towards a hyperlinked society: A critical review of link studies. *New Media & Society*, *15*(5), 737–751. doi:10.1177/1461444812462851

De Maeyer, J., & Holton, A. E. (2016). Why linking matters: A metajournalistic discourse analysis. *Journalism*, *17*(6), 776–794. doi:10.1177/1464884915579330

De Maeyer, J., & Le Cam, F. (2015). The material traces of journalism. *Digital Journalism*, *3*(1), 85–100. doi:10.1080/21670811.2014.928021

Deuze, M. (2007). *Media work*. Cambridge, UK: Polity.

Dimitrova, D. V., Connolly-Ahern, C., Williams, A. P., Kaid, L. L., & Reid, A. (2003). Hyperlinking as gatekeeping: Online newspaper coverage of the execution of an American terrorist. *Journalism Studies*, *4*(3), 401–414. doi:10.1080/14616700306488

Editor & Publisher Directories. (2006). *Editor & Publisher international yearbook*. New York, NY: Author.

Etling, B., Kelly, J., Faris, R., & Palfrey, J. (2009). *Mapping the Arabic blogosphere: Politics, culture and dissent* (Berkman Center Research Publication No. 2009-06). Cambridge, MA: Berkman Center for Internet & Society.

Gaines, C. (2016, March 14). News update: Facebook Instant Articles to be unleashed [Blog post]. Retrieved from https://www.socialmediadelivered.com/blog/2016/03/14/news-update-facebook-instant-articles

Gao, Y., & Vaughn, L. (2006). Web hyperlink profiles of news sites: A comparison of newspapers of USA, Canada and China. *ASLIB Proceedings, 57*(5), 398–411.

Gilbert, C. G. (2005). Unbundling the structure of inertia: Resource versus routine rigidity. *Academy of management journal, 48*(5), 741–763. doi: 10.5465/AMJ.2005.18803920

Gonzales-Bailon, S. (2009). Opening the black box of link formation: Social factors underlying the structure of the Web. *Social Networks, 31*(4), 271–280. doi:10.1016/j.socnet.2009.07.003

Greer, J. D., & Mensing, D. (2004). U.S. news Web sites better, but small papers still lag. *Newspaper Research Journal, 25*(2), 98–112. doi:10.1177/073953290402500207

Greer, J. D., & Mensing, D. (2006). The evolution of online newspapers: A longitudinal content analysis, 1997–2003. In X. Li (Ed.), *Internet newspapers: The making of a mainstream medium* (pp. 13–32). Mahwah, NJ: Erlbaum.

Grotke, A. (2011). *Collecting and preserving user-generated content: Web archiving at the Library of Congress*. Paper presented at the Computers in Libraries Conference, Washington, DC. Retrieved from http://www.infotoday.com/CIL2011/Presentations.asp

Gunter, B. (2003). *News and the Net*. Mahwah, NJ: Erlbaum.

Halavais, A. (2008). The hyperlink as an organizing principle. In J. Turow & L. Tsui (Eds.), *The hyperlinked society: Questioning connections in the digital age* (pp. 39–55). Ann Arbor, MI: University of Michigan Press.

Heiman, A., & Clarysse, B. (2005). The imprinting effect of initial resources and market strategy on the early growth path of start-ups. *Academy of Management Proceedings, 2005*(1), A1–A6. doi:10.5465/AMBPP.2005.18778595

Hsu, D. H., & Lim, K. (2013). Knowledge brokering and organizational innovation: Founder imprinting effects. *Organization Science, 25*(4), 1134–1153. doi:10.1287/orsc.2013.0863

Huber, G. (1991). Organizational learning: The contributing processes and the literatures. *Organization Science, 2*(1), 88–115. doi:10.1287/orsc.2.1.88

Johnson, V. (2007). What is organizational imprinting? Cultural entrepreneurship in the founding of the Paris Opera. *American Journal of Sociology, 113*(1), 97–127. doi:10.1086/517899

Ketterer, S. (2001). Links engage readers of online crime stories. *Newspaper Research Journal, 22*(2), 2–13. doi:10.1177/073953290102200201

Kim, Y. S., Yang, Y. S., & Hwang, H. (2002). Structural expansion and the cost of global isomorphism: A cross-national study of ministerial structure, 1950–1990. *International Sociology, 17*(4), 481–503. doi:10.1177/0268580902017004002

Kleinberg, J. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM, 46*(5), 604–632. doi:10.1145/324133.324140

Kogut, B. (1993). Learning, or the importance of being inert: Country imprinting and international competition. In S. Ghoshal & E. Westney (Eds.), *Organization theory and the multinational corporation* (pp. 136–154). London, UK: Palgrave Macmillan.

Lee, A. M. (1937). *The daily newspaper in America*. New York, NY: Macmillan.

Leung, A., Foo, M. D., & Chaturvedi, S. (2013). Imprinting effects of founding core teams on HR values in new ventures. *Entrepreneurship Theory and Practice, 37*(1), 87–106. doi:10.1111/j.1540-6520.2012.00532.x

Li, X. (2006). *Internet newspapers: The making of a mainstream medium*. Mahwah, NJ: Erlbaum.

Lim, J. (2012). Power relations among popular news websites for posting headlines through monitoring and imitation. *New Media & Society*. Advance online publication. doi:10.1177/1461444812466716

Marquis, C. (2003). The pressure of the past: Network imprinting in intercorporate communities. *Administrative Science Quarterly, 48*(4), 655–689. doi:10.2307/3556640

Marquis, C., & Tilcsik, A. (2013). Imprinting: Toward a multilevel theory. *The Academy of Management Annals, 7*(1), 195–245. doi:10.1080/19416520.2013.766076

Massey, B. L., & Levy, M. R. (1999). Interactivity, online journalism, and English-language Web newspapers in Asia. *Journalism & Mass Communication Quarterly, 76*(1), 138–151. doi:10.1177/107769909907600110

McManus, J. H. (1994). *Market-driven journalism: Let the citizen beware?* Thousand Oaks, CA: SAGE Publications.

Mecca, J. T., & Mumford, M. D. (2014). Imitation and creativity: Beneficial effects of propulsion strategies and specificity. *The Journal of Creative Behavior, 48*(3), 209–236. doi:10.1002/jocb.49

Meraz, S. (2011). Using time series analysis to measure intermedia agenda-setting influence in traditional media and political blog networks. *Journalism & Mass Communication Quarterly, 88*(1), 176–194. doi:10.1177/107769901108800110

Milanov, H., & Fernhaber, S. A. (2009). The impact of early imprinting on the evolution of new venture networks. *Journal of Business Venturing*, *24*(1), 46–61. doi:10.1016/j.jbusvent.2007.11.001

Ognyanova, K., & Monge, P. (2013). A multitheoretical, multilevel, multidimensional network model of the media system: Production, content, and audiences. *Annals of the International Communication Association*, *37*, 66–93. doi:10.1080/23808985.2013.11679146

Park, H. W., Kim, C. S., & Barnett, G. A. (2004). Socio-communicational structure among political actors on the Web in South Korea: The dynamics of digital presence in cyberspace. *New Media & Society*, *6*(3), 403–423. doi:10.1177/1461444804042522

Park, H. W., & Thelwall, M. (2003). Hyperlink analyses of the World Wide Web: A review. *Journal of Computer-Mediated Communication*, *8*(4). doi:10.1111/j.1083-6101.2003.tb00223.x

Perkmann, M., & Spicer, A. (2014). How emerging organizations take form: The role of imprinting and values in organizational bricolage. *Organization Science*, *25*(6), 1785–1806. doi:10.1287/orsc.2014.0916

R Core Team. (2015). R: A language and environment for statistical computing (Version 3.1.3) [Computer software]. Retrieved from https://www.r-project.org/

RStudio. (2015). RStudio: Integrated development environment for R (Version 0.99.441) [Computer software]. Retrieved from https://www.rstudio.com/products/rstudio/features/

Schwarzlose, R. (1987). *Newspapers: A reference guide*. New York, NY: Greenwood Press.

Scott, B. (2005). A contemporary history of digital journalism. *Television & New Media*, *6*(1), 89–126. doi:10.1177/1527476403255824

Shinkle, G. A., & Kriauciunas, A. P. (2012). The impact of current and founding institutions on strength of competitive aspirations in transition economies. *Strategic Management Journal*, *33*(4), 448–458. doi:10.1002/smj.1946

Shumate, M. (2012). The evolution of the HIV/AIDS NGO hyperlink network. *Journal of Computer-Mediated Communication*, *17*(2), 120–134. doi:10.1111/j.1083-6101.2011.01569.x

Shumate, M., & Lipp, J. (2008). Connective collective action online: An examination of -the hyperlink network structure of an NGO issue network. *Journal of Computer-Mediated Communication*, *14*, 178–201. doi:10.1111/j.1083-6101.2008.01436.x

Sydow, J., Schreyögg, G., & Koch, J. (2009). Organizational path dependence: Opening the black box. *Academy of Management Review*, *34*(4), 689–709.

Thompson, D. R., & Wassmuth, B. L. (2001). Few newspapers use online classified interactive features. *Newspaper Research Journal*, *22*(4), 16. doi:10.1177/073953290102200403

Tremayne, M. (2004). The Web of context: Applying network theory to the use of hyperlinks in journalism on the Web. *Journalism & Mass Communication Quarterly*, *81*(2), 237. doi:10.1177/107769900408100202

Tremayne, M. (2005). News websites as gated cybercommunities. *Convergence*, *11*(3), 28–39. doi:10.1177/135485650501100303

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge, UK: Cambridge University Press.

Weber, M. S. (2012). Newspapers and the long-term implications of hyperlinking. *Journal of Computer-Mediated Communication*, *17*(2), 187–201. doi:10.1111/j.1083-6101.2011.01563.x

Weber, M. S., & Monge, P. (2011). The flow of digital news in a network of authorities, hubs and providers. *Journal of Communication*, *61*(6), 1062–1081. doi:10.1111/j.1460-2466.2011.01596.x

Weber, M. S., & Nguyen, H. (2015). *Big data? Big issues: Degradation in longitudinal data and implications for social sciences*. Paper presented at the WebSci 2015, Oxford, UK.