Matthew S. Weber

The tumultuous history of news on the web

2017

https://doi.org/10.25969/mediarep/12518

Veröffentlichungsversion / published version
Sammelbandbeitrag / collection article

Empfohlene Zitierung / Suggested Citation:
Weber, Matthew S.: The tumultuous history of news on the web. In: Niels Brügger, Ralph Schroeder (Hg.): *The Web as History. Using Web Archives to Understand the Past and the Present.* London: UCL Press 2017, S. 83–100. DOI: https://doi.org/10.25969/mediarep/12518.

4

The tumultuous history of news on the web

Matthew S. Weber

News and the web

Tuchman (1978) noted in her ethnographic research that newspapers were responsible for creating a constant flow of information to consumers, continually moulding our comprehension of society. In the 1970s, more than 62 million newspapers were sold in the USA each day. Readership had been growing or stable for more than 50 years. Subsequently, what was a constant flow has become a torrent of news today, and the news industry has entered a period of remarkable tumult. News and information today flows to consumers via many traditional media, but it is increasingly complemented, and in some cases preceded, by computers, tablets, mobile phones and other emerging devices. The internet today is the pipeline that feeds information to consumers, and for the time being the web is a primary window through which that information is distributed, accessed and retrieved.

In the past 20 years, the relationship between news media companies and internet technology has been a tenuous one at best. Traditional print newspaper circulation has declined sharply since the early 1990s; hedging their bets, many newspaper companies began experimenting with internet technology as early as the late 1980s. Much of that experimentation, however, was often done simply as an extension of existing printed products (Boczkowski, 2004a). Most newspaper executives were sceptical that the web represented a credible threat, and as a result most innovation occurred at arm's length (Chung, 2007). In recent years, however, the need for change has been widespread and notable; even television news has faced challenges from the rise of online video

and the looming threat of online-only programming (Perren, 2010). Thus, although there were early innovators in the newspaper industry, it is only in recent years that revolutionary change has occurred on a large scale (Karimi and Walter, 2015; Schlesinger and Doyle, 2015).

History of news on the web, from the web

The following discussion traces the tumultuous history of news media on the web. More specifically, this chapter focuses on the history of newspapers in the USA as they have grappled with adapting to new digital technology by tracing their development through their websites. A multitude of notable volumes exist that trace change in the newspaper industry in response to digital technology (see, for example, Boczkowski, 1999, 2004a; Kawamoto, 2003; Usher, 2014). Contrasting prior work, this chapter is not intended as a comprehensive history of news online; rather, this chapter aims to illustrate the adaptation of the newspaper industry to the web through an examination of the content of the web itself, including snapshots of the broad ecosystem.

The web as history: news online

This story is told through an examination of archived news media content maintained within the Internet Archive. There are few cohesive sources of archived news content available for large-scale research; the largest source of archived internet content is available via the Internet Archive (accessible via archive.org). Founded in 1996, the San Francisco-based Internet Archive is a non-profit organization that was established with a mission to preserve the history of the web, and to build an internet library containing that history. The Internet Archive is best known for the Wayback Machine. The Wayback Machine is the graphic interface for the Internet Archive databases, and allows users to freely access the information stored within the archive. Beyond the Wayback Machine, the Internet Archive maintains a rich repository of web pages, photographs, videos and other types of digital content. As of 2015, the Internet Archive had recorded, parsed and archived 438 billion web pages occupying 23 petabytes (PB) of storage space. While the Internet Archive does not contain a complete record of the internet, it is the single most extensive archive of the internet.

Most of the statistical analysis and web pages referenced in following sections were enabled through the ArchiveHub project.

ArchiveHub is a National Science Foundation-funded project intended to enable researcher access to archival internet content. The ArchiveHub project includes a substantial collection of media content stored in a researcher accessible database format, with extensive metadata to aid scholarly work. For example, a dataset of media websites from 2008 to 2013 includes 1.3 billion captures of web pages representing 540 million unique URLs. A second smaller dataset contains a total of 25,628 websites from 1996 through 2007; the dataset contains roughly 300 million total captures. The aggregate datasets provide a robust source for better understanding the history of news and newspapers on the web. The focus of the ArchiveHub project is on the development of tools that extract hyperlinks from Internet Archive records. Hyperlinks are useful for understanding the degree of connectivity between websites, and for mapping the flow of information (Weber and Monge, 2011). In addition to extracting hyperlinks, the tools provide additional information including keywords relevant to a given web page, and information such as the size of the archived website content measured in megabytes (MB).

Thus, in the context of this chapter, the web is both the source of change, and the means by which the story is told. The dramatic change that has occurred in the news media industry was predicated by the introduction of new technological standards in the 1990s; and yet today I am able to tell that story thanks to that very technology. Recent scholarship examining online news has often relied on data from the web to help tell the story of changes in the modern news media landscape. A 2012 study examined newspaper websites in the UK. This illustrated the fluidity of online news, with sources and content being added and deleted over time as a particular story or news event evolved; the traditional notion of a news article as a fixed unit is less prominent in the digital space (Saltzis, 2012). Similarly, web content has been used to show how competitive news outlets often mimic the coverage of competitors in order to ensure relevance (Boczkowski, 2010).

I draw on a number of different examples and analyses to illustrate the role of the web in telling the history of news on the web. Large-scale data are paired with a series of case studies to demonstrate how individual news media responded to competitive threats, first from hyperlinks and the free flow of content, and subsequently from social media. The growth of internet technology flattened the competitive landscape for newspapers; while much has been made of innovation within industries, the interaction between print newspapers and new media brought about rapid transformation across the news landscape. The following

vignettes focus first on the early period of the web, from 1990 through 2005; the latter sections fast-forward and highlight more recent developments between 2010 and 2015.

The early days of news on the web: 1990–2005

In 1991, Tim Berners-Lee published the first website on the first web server at the CERN laboratory in Switzerland (Berners-Lee, 1991). Newspapers were quick to join the web, building on their early experimentation with internet technology. The introduction of new technology delivered a shock to the newspaper population, but the industry responded with measured innovation rather than any significant restructuring (Sylvie and Witherspoon, 2002; Boczkowski, 2004a, 2004b; Patterson, 2007). In this way, newspapers have taken advantage of internet technology since its infancy, although many early experiments were not successful. The Columbus Dispatch was the first daily newspaper in the USA to provide an online version for its customers. In 1980 the newspaper provided access to an online version via the internet service provider CompuServe (Kawamoto, 2003). A number of newspapers also experimented with Videotex, an early digital information transmission system. Knight-Ridder, a former American news company, even developed a proprietary Videotex system.

The newspaper industry, in general, viewed online technology as a new medium for distributing an existing product, and for nearly a decade newspaper organizations focused on products that simply delivered the print product digitally (Boczkowski, 2004a). To this end, Falkenberg (2010) delineates between 'online newspapers' and 'newspapers on the web', with the early period of the web primarily occupied by newspapers on the web as replications of their printed products. In 1993, the first commercial graphic web browser, Mosaic, was launched, and by 1999 more than 4,900 newspapers globally had launched web versions of their newspapers.

Web technology gave rise to the newspaper websites that many consumers are familiar with today; 1991 to 1993 represented a juncture in the history of online news, because World Wide Web protocols including hypertext markup language (HTML) enabled a new visual interface for accessing news via the internet (Stovall, 2004). In 1994 *Raleigh News & Observer* launched Nando.net as a web-based version of their newspaper; this is one of the first examples of a web-based newspaper living on the web outside of an internet service provider's intranets. The first

available record of Nando.net is available in the Internet Archive. The visual nature of the interface is clear, as is the differentiation from the traditional print product.

The archival pages of Nando.net include articles, images and hyperlinks to other early websites. This type of record allows for an examination of the type of content available on the early web, as well as the type of communication enabled by early web protocols. For instance, the earliest records of Nando.net include a rich repository of photographs drawn from news wire services, contrary to the perception that early web-based newspapers lacked graphics or visuals. The *Raleigh News & Observer* generally published photographs with sports content, but photographs were for the most part not included with general news articles. Based on statistics on archived web pages of Nando.net, the website received about 14 million visitors per week in 1999. Despite the arguable success of early web ventures such as Nando.net, newspapers were quickly failing to replace lost print advertising revenue with the equivalent in online advertising revenue (Weber and Monge, 2014).

The rise of blogs

Newspaper advertising revenue has generally been tied directly to the number of readers purchasing a newspaper. In the 1990s, new digital sources were drawing readers away from printed products in large numbers, and in time, rapidly drawing away revenue.

In 1997 and 1998, the first variations of a new type of web content emerged online in the form of weblogs, more broadly referred to as blogs. The World Wide Web was not a tool born of the newspaper industry. Rather, it is a technological disruption that originated as a government and research communication tool. Consequently, innovation on the web came from many sources. By 1994, early innovators were using web pages as tools for online diaries and personal commentaries. In 1997, Jorn Barger launched Robot Wisdom, which featured a listing of links that Barger liked to visit, as well as updates from Barger's daily life (Rettberg, 2008). Similar types of sites began to pop up en masse, but generally failed to attract large audiences. The term 'blog' was first used to describe these sites by Peter Merholz in 1999; the term was a shortening of 'weblog' (web log), which Merholz thought ought to be pronounced 'wee-blog' and later shortened to 'blog'. The large-scale emergence of blogs served as an exchange network of sorts, whereby users were able to share hyperlinks with one another to identify information sources of interest (Ammann, 2011). But a high barrier to entry

plagued these early sites; a user seeking to build a blog was required to have a sufficient amount of technological expertise in order to build and maintain the site. In October 1998, however, Open Diary was founded to offer users space on the web with free hosting and easy-to-use online publishing options. Within four months, the site had 25,000 hosted online diaries. Pitas launched in 1999 offering free blogging tools, followed by the launch of Blogger. Traditional newspapers continued to adhere to the strict routines of printed newspapers, but blogs allow writers and reporters to share opinions and publish relatively raw content outside the bounds of journalistic hierarchy. Early blogs were relatively simple hypertext documents updated on a relatively frequent basis, with content ranging from a few roughly assembled sentences to complete magazine-length features (Matheson, 2004).

Weblogs were one of the first forms to dramatically reinvent the form of daily news. The Drudge Report, an early variant of weblogs, first appears in the Internet Archive on 10 December 1997 (the Drudge Report was founded in 1996, but in the early archive there is an occasional delay between actual founding and appearance in the database). The Drudge Report and other 'news' blogs were simply aggregations of links to other websites and news articles. The Drudge Report continues as a popular and successful news source today. Early blogs were a harbinger of future change, but were rudimentary in nature.

And then came social media

In 1997, SixDegrees.com launched, allowing users to create profiles and connect with other friends on the site. SixDegrees is generally credited with being the first social networking site (SNS) (boyd and Ellison, 2008). Subsequently, numerous imitators emerged, and many were successful in improving social networking as a platform. In 2002, Friendster.com launched and quickly gathered a following of more than 300,000 users. From 2003 onward, SNSs established themselves as mainstream media platforms, due largely to the development of Web 2.0 technology. Web 2.0[1] technologies are a class of platforms that enable consumer participation and interaction in online environments, including discussion and creation of the news. Today, an SNS is viewed as a website that 'connects and presents people based on information gathered about them, as stored in their user profile' (Cruz-Cunha et al., 2011: xviiii). boyd and Ellison (2008) distinguish SNSs as websites that allow users to (1) create a public-facing profile, (2) construct a list of users to whom they are connected, and (3) navigate lists of connections

for individuals and their connections. More broadly, SNSs are online resources that allow users to create 'maps' of their social networks, and to share information through these networks. By 2000, SNSs gained further traction as a means of sharing information between users; eventually this included pointing others to news articles, and providing links to news on websites (Suler, 2004).

The challenge of adapting to the web

The introduction of World Wide Web protocols was a first tipping point; the rise of Web 2.0 technology created a second tipping point in the history of news on the web. By the turn of the century, social networking sites were gaining in number and popularity. With the popularization of Web 2.0 technology, blogs became increasingly widespread and interactive. During this period, successful news blogs such as Huffington Post and Gawker were launched. Thus, during this period the notion of online newspapers began to reach maturity, as content was being produced exclusively for the web (Falkenberg, 2010). For instance, when the Huffington Post officially launched on 9 May 2005, its interface was driven by a strong visual design, and the site included features that allowed users to comment on the news and to engage with the website.

The rise of bloggers and blogs presented a clear challenge to traditional newspapers, as did the rise of Web 2.0 content and social media. In part, newspapers were challenged by their own structural inertia; given the storied history of many of the large newspapers in the USA, it is not surprising that many organizations were hesitant to transform completely to a digital platform (Weber, 2012; Weber and Monge, 2014).

For example, Figures 4.1 and 4.2 illustrate the changing dynamics of interaction between traditional newspapers on the web, and blogs on the web. The red circles represent established newspapers with a presence on the web; the blue circles represent blogs, online communities, and online-only news sources. The data shown in the illustrations focuses on a subset of the larger dataset, with the subset containing 269 blogs, 192 online communities and social networking sites and 487 newspapers. Web archiving technology changed significantly during this early period, and thus the subset was selected by identifying the websites for which data was consistently available for the period of interest. Hyperlinks to advertising websites were removed. A connection between two websites exists if a hyperlink existed between two websites and was present at least three times in a given year. Hyperlinks are useful for analysing the relationship that existed between media

organizations; in cases where a hyperlink persists over time, prior research has established a connection between the presence of hyperlinks, and the presence of a relationship between organizations (Gao and Vaughn, 2006; Shumate and Lipp, 2008; Tsui, 2008; Turow and Tsui, 2008).

As Figures 4.1 and 4.2 demonstrate, there was little interaction between bloggers and newspapers during this early period. There is a clear shift over time, however, as the two disparate groups become more intertwined. The visualizations were generated by conducting an analysis of the hyperlinks between newspaper websites and bloggers via the Internet Archive, and illustrate the connections between all websites in the network.

A subsequent statistical analysis provides further insight. For instance, the density of the networks visualized in Figures 4.1 and 4.2 measures the percentage of hyperlinks that exist as a percentage of the total possible hyperlinks. Controlling for growth of the web as a whole,

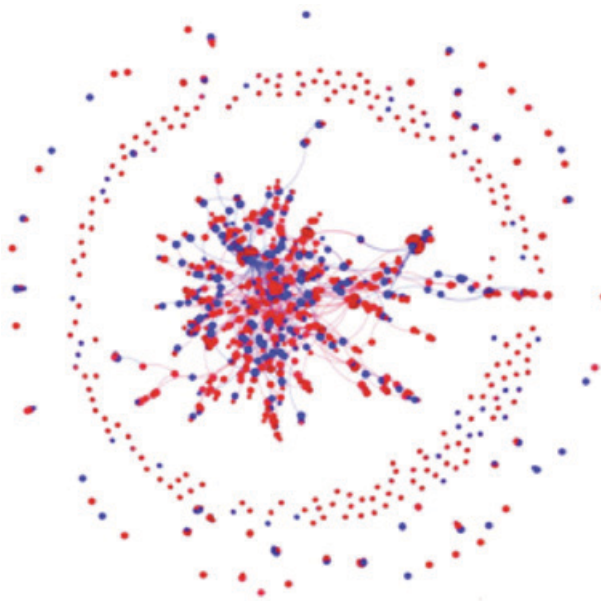

Figure 4.1 Connections between newspapers and other websites on the web in 1999 (red indicates traditional newspapers; blue indicates online native entities)

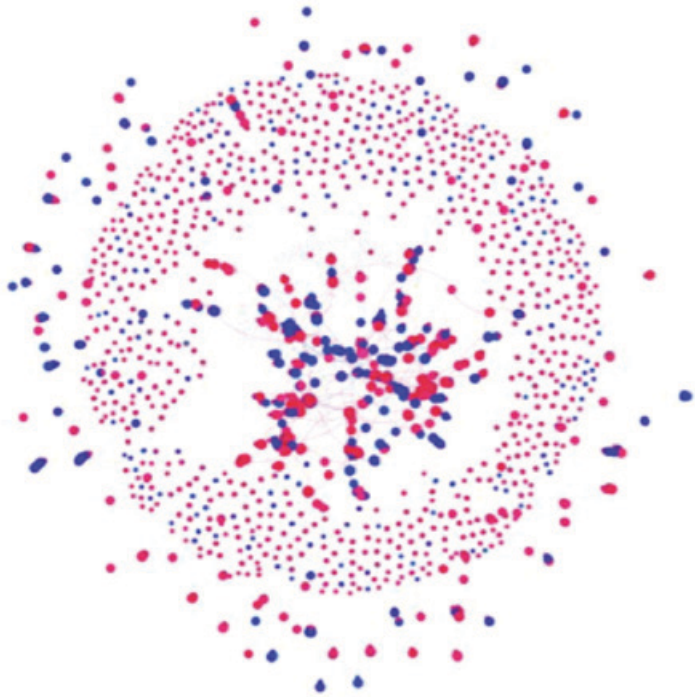

Figure 4.2 Connections between newspapers and other websites on the web in 2005 (red indicates traditional newspapers; blue indicates online native entities)

the density of the network decreased; in 1999 the graph density was 0.34%, and in 2005 it had decreased to 0.22%. The web has always been a vast space, and the increase in news outlets and websites over time helped to create 'pockets' of news early on. Websites clustered together in relatively disparate groups, as evidenced by the interconnectivity of early websites. At the same time, connections between different types of websites increased; in 1999, only 12% of hyperlinks in the above sample existed between website types; that number increased to 32% in 2005. This helps to explain why the visualization is more cohesive in the later time period, and demonstrates the slow erosion of the barriers between the disparate pockets of news.

By the end of this critical period, the online news ecosystem no longer existed as an ecosystem of disparate entities. The new news

ecosystem that existed in 2005 developed as an agglomeration of traditional newspapers, blogs and early social networking sites; many of the virtual barriers were eroded by this point in time. The early period of transformation saw the online news ecosystem move from one of isolated hubs of news websites – dispersed with relatively little interconnectivity – to an integrated network of news and information websites.

The web, grown up: 2010–2015

What a difference a decade makes. According to the Pew Research Center's 2012 News Consumption Survey (Kohut et al., 2012), in 2000, 23% of Americans reported that they went online for news at least three days a week. In 2010, that number had increased to 46%. From 2010 onwards, the disruption of the news media landscape accelerated. This is, in part, due to broad societal shifts. Based on data from the US Census, the percentage of households with internet access increased from 41.5% in 2010 to 71.1% in 2015. The number of web-based sources increased with the growth in online traffic, and yet, users were increasingly concentrated among a handful of websites. In 2010, an analysis of 4,600 news and information websites tracked by Nielsen Media found that the top 7% of news media websites collected 80% of the overall traffic (Mitchell and Rosenstiel, 2010). According to ComScore ('Digital: Top 50 online news entities', 2015), by 2015, the top online sources for news in the USA were Yahoo!/ABC News Network, CNN, NBC News Digital, HuffingtonPost.com and CBS News; the top newspaper, *USA Today*, came in at sixth in terms of unique visitors.

During this period, newspaper print circulation continued to slip, although the rate of decline slowed to less than a 5% decline per year. Yet significant shifts in audience preferences had already occurred. Data from the American Press Institute found in 2014 that 69% of Americans accessed news via their computers, 56% accessed news via their cell phones and 29% accessed via their tablets (based on sources used for news in the past week). Audience preferences had clearly taken a dramatic swing towards a wide array of digital devices.

Revenue challenges persisted hand-in-hand with readership challenges; digital revenue continues to be a small fraction of overall revenue for legacy newspapers, as well as other legacy news sources including magazine and television. Despite the slow decline of print newspapers, many continued to attract a substantial audience. In print, *USA Today* had an average daily print circulation of 3.3 million, based on 2014 data.

Comparatively, however, *USA Today* is estimated as the largest online newspaper in the USA, reaching 54.5 million users online in January 2015. Elsewhere, *The New York Times* maintained a daily circulation of 2.1 million in 2014. *The New York Times* has also had success driving revenue online; the newspaper is an example of a traditional news organization that has succeeded in developing a revenue model for its online content (Pickard and Williams, 2013); despite this success, a relatively small proportion of print newspapers have succeeded in replicating a 'paywall' model.[2]

At the same time, native online news services began to emerge as key providers of news and information; SNS are increasingly driving consumers to news content, and serving as a key portal through which consumers discover news (Perelman, 2014). Native online news services refer to organizations that create and distribute news media solely through the web. Today, there are many different iterations of native online news services; despite the popularity of online news platforms, there is still significant experimentation. This is consistent with a growing industry (Weber and Monge, 2014), and is a trend that will likely continue for the near-term future. For example, Buzzfeed first gained prominence when it launched in 2006 as an incubator of digital content, but as it has evolved into a news provider it has gone on to develop its own newsroom (LaFrance and Meyer, 2015). In another example, ProPublica was launched in 2007 as a non-profit news organization. The news service's primary goal is to publish in-depth investigative journalism; in 2010, ProPublica became the first online news service to win a Pulitzer Prize. But not every venture has been successful; for instance, GigaOm, a popular technology news blog, launched in 2006 but shut down in 2015 due to declining revenue.

The changing news landscape as a local story

As the preceding discussion illustrates, the period of time from 2010 to 2015 has seen continued changes in the news media landscape. Despite ominous predictions about the death of the printed newspaper, and despite the contraction of the newspaper industry as a whole (Deuze, 2003), newspapers continue to publish both in print and online, albeit in a diminished capacity. More recent data from the Internet Archive, covering 2010 to 2015, provides a snapshot of recent changes to the national news landscape. For instance, an examination of outbound hyperlinking of the top 25 national newspapers in the USA reveals that

from 2012 onwards, 98% of those newspapers' websites contained outbound links to Twitter and Facebook, illustrating the growing role of SNS in the news ecosystem.

As the web has grown up, one area where there has been a profound change is in the provision of local news. Much of the early innovation by newspapers occurred at larger newspapers; small community newspapers lagged behind in terms of the development of web content (Greer and Mensing, 2004). And yet, community engagement with local news is well established as a key predictor of community health, and helps to foster social interaction within communities (Paek et al., 2005). Readership of local news is also directly related to the likelihood of voting in elections (Moy et al., 2004).

As web-based news moved towards maturity, many saw an opportunity for web-platforms to improve local news coverage. For instance, Downie and Schudson (2009) predicted that the launch of local news websites by entrepreneurial journalists would help to improve local democracy. Similarly, Lewis (2011) observed that foundation-funded hyperlocal websites managed to successfully pair new technology with high quality journalism, creating an opportunity for growth.

Despite early optimism, by most accounts the digitization of news has had a negative impact on local news. As of 2010, many local newspapers were in a state of crisis (Nielsen, 2015; Wadbring & Bergström, 2015), and in markets where coverage had decreased there was already an indication of a decline in political participation (Hayes and Lawless, 2015). A 2015 report from the Democracy Fund highlights the plight of local news; the report examined the health of local news in New Jersey, and found that there are stark gaps in local news coverage (Napoli et al., 2015). For instance, the report observed that there are only 0.58 sources of news for every 10,000 people in Newark, NJ, with a population of 277,000 and a per capita income of $13,009. Comparatively, Morristown, NJ, with a population of 18,000 and per capita income of $37,573, has 6.11 sources per 10,000 people.

A new perspective on change in local news coverage is enabled through an examination of the Internet Archive's records of local news websites. In order to better assess changes in local news ecosystems, a longitudinal examination of the New Jersey news ecosystem was conducted using a subset of data from the Internet Archive. The subset of websites was selected by hand coding local news websites in New Jersey and extracting those websites from the Internet Archive's repositories. In this case, an analysis of the hyperlinks between local news websites allows for an examination of both the scale of the local news ecosystem,

as well as the cohesiveness and coherence of the ecosystem. The analysis is based on a subset of local news websites extracted from the Internet Archive. A list of 390 local New Jersey news websites was created based on websites that operated between 2008 and 2012; because of the structure of the market, some key websites from Philadelphia and New York were included as key information sources. The resulting dataset includes approximately 1.6 million captured websites across the five-year period, which includes both the focal local news websites as well as websites that are connected to those organizations. Moreover, the number of captures does not directly reflect the number of websites. It is not unusual for a website to have hundreds, or even thousands, of web pages with a given domain, and those domains are crawled many times in a given year.

Again, using the ArchiveHub system, hyperlinks between websites were extracted from the Internet Archive data. This allowed for an examination of the flow of information between news websites (see Weber, 2012, for a further discussion of the role of hyperlinks as a tool for guiding information flow). In addition, it was possible to summarize the amount of information within each domain (measured in megabytes). The amount of information in megabytes can be considered a proxy for the amount of text and images on a web page.

Table 4.1 provides descriptive information regarding the state of local news in New Jersey from 2008 through 2012. In order to examine changes in the local news landscape, this analysis focuses on changes in the core sample of websites, based on average degree, average path, density, connected strong components and clustering. Average degree measures the average number of connections per website. Average path measures the average of the shortest path that exists between all

Table 4.1 Network analysis of local New Jersey news websites, 2008–2012

	2008	2009	2010	2011	2012
Websites	90	193	216	338	203
Connections	105	246	261	388	315
Average degree	1.17	1.28	1.24	1.14	1.56
Average path	1.48	2.25	2.05	3.11	2.45
Density	0.013	0.007	0.006	0.003	0.008
Connected strong component	90	106	207	335	194
Clustering	0.014	0.029	0.0016	0.007	0.095

websites; it gives an indicator of how connected all websites are to one another. Density, as previously mentioned, accounts for the number of connections between websites as a percentage of the possible connections. The number of connected strong components gives a measure of the number of clusters of websites that exist. Clustering gives a measure of the degree of clustering on a scale of 0 to 1.

The descriptive analysis in Table 4.1 reveals some critical changes in the local New Jersey news landscape. Between 2008 and 2011, the number of local news websites increased significantly, but the data begin to reflect the broader trend of decline in 2012. Although the number of websites present declines in 2012, the connectedness of websites remains relatively stable, as seen in the number of connections, as well as changes in the average degree and path length. With regards to clustering, the number of components increased at first, but declined by 2012. On the other hand, the degree of clustering decreased.

In aggregate, these data illustrate a story of a decreasing number of local news websites that are increasingly clustered together. Echoing prior research, this type of story would be consistent with a declining number of websites increasingly sharing information with one another.

Figures 4.3 and 4.4 provide further context, illustrating the connections that existed between key websites in the New Jersey local news ecosystem based on hyperlinking. The two visualizations illustrate the top 30% of websites in each year, based on the degree of connectivity to other websites.

Comparing the two visuals, it is clear that the ecosystem in 2012 had become more tightly clustered. Moreover, there are fewer organizations engaged in this central cluster; for instance, Philly.com is no longer central in this network, and the dailyrecord.com is less prominent. Looking to 2012, the connections within this smaller cluster are also stronger, as illustrated by the thickness of the connections between websites. Orange connections represent relationships with less prominent websites, whereas blue connections are relationships between equally popular websites.

This analysis provides an overview of the changing local news landscape at the beginning of this critical period. Clearly this is a single analysis of a single state, but the trends are consistent with previous research on the topic. Large-scale web data provides a unique vantage point for assessing the health of local news environments; furthermore, the nature of web archives provides a tool with which research can code and analyse the actual content, creating a fertile resource for future research.

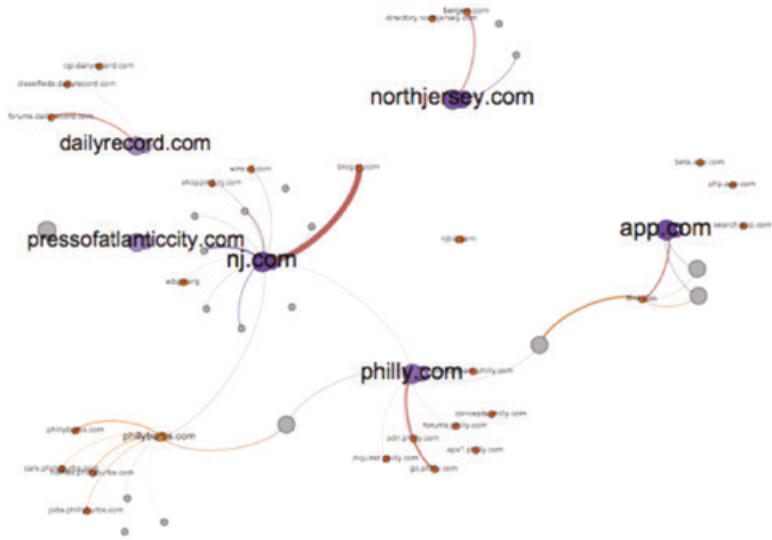

Figure 4.3 New Jersey local news ecosystem, 2008

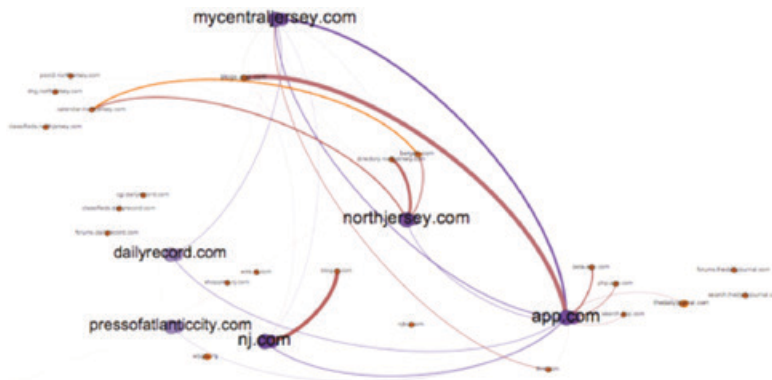

Figure 4.4 New Jersey local news ecosystem, 2012

The next generation of news on the web, and beyond

In sum, despite the promise of digital news platforms, traditional newspaper models have struggled in a web-based environment. National newspapers have transitioned to digital platforms, but success is largely isolated among the few largest newspapers. Local newspapers, on the other hand, are continuing to decline, and local communities are suffering as a consequence.

In aggregate, prior research, and the illustrations provided in this chapter, underscore the ongoing turmoil within the news media industry. The continuing growth of new platforms shows that change is expected to continue, and in many cases, news is moving away from the web towards mobile and application based platforms.

Echoes of social networking

The role of the web is perhaps as unclear today as it was in 1995; this is particularly true in light of recent trends. For instance, today SNSs are mainstream sources of information. In 2015 Facebook reported that it had 1 billion active users in a single day. SNSs are shifting the way that they provide consumers with news. In recent years, SNSs have been a key driver of consumers to newspaper websites. For instance, a 2014 Pew Report found that 30% of the general population cited Facebook as the place they turned to for news; in turn, those users are then connected via hyperlinks to specific newspaper websites (Anderson and Caumont, 2014).

In 2015, moreover, a number of SNSs have developed new platforms that provide consumers with news from content partners within the SNSs' own application. SNSs have quickly developed into platforms where users can discover and engage with news (Winter et al., 2015), and it is increasingly unnecessary to go to other websites. Thus, distributed content networks that draw content from partners onto SNSs keeps users on sites such as Facebook, Twitter or Snapchat, and drives revenue for content providers through in-application advertising (Benton, 2015). Facebook introduced its Instant Articles platform, and Apple introduced its Apple News application; both provide content from third-party providers directly in the application. Through partnership agreements, some of the applications also provide consumers with exclusive content. For example, Apple formed a partnership with Wired Magazine that included exclusive access to certain feature news articles. SNSs are quickly shifting to control the distribution of news media, and moving to a system whereby news media is distributed by a given platform's own application.

Mobility, automation and everything after

The growth of SNSs as a platform for news, and the increasing prominence of mobile applications, further underscores a move away from the distribution of content solely via the web. For instance, of the

54.5 million users who accessed USA Today online in January 2015, 34 million of those users accessed the website via mobile devices. In 2012, 39% of survey respondents indicated they received their news in the past 24 hours from a mobile device (Kohut et al., 2012). Mobile platforms often access information via the web, but are increasingly reliant on applications. The growth of media applications has created yet another new avenue for news distribution. In another vein, the move to digital news on the web has increased the speed with which consumers receive information, but future iterations are likely to see the provision of news automated to a certain degree. Indeed, the move to mobility and application-based news is also bringing forth a new focus on algorithms and the automation of information flow. Algorithms are increasingly being used to automatically decide what news consumers see based on their reading habits (Mysiani, 2013). Others are working to develop algorithms that produce summaries of content and routine news stories such as financial summaries and sports recaps, automating the actual production of the news (Sood et al., 2007; Liu and Birnbaum, 2008).

Why the web's history of news matters

In the midst of previous and forthcoming change, the web provides a critical resource for examining the nature of change for newspapers. Not only does the history of the web allow for sensemaking of previous changes, but it also provides context for understanding future change. As a history of news, the web also provides a critical record for understanding what actually was. For example, archived internet records provide one of the best records of the *Rocky Mountain News*, a daily Denver newspaper that ceased publication in 2009. The history of the web can even provide a critical perspective on news events.

For example, on 1 May 2003, then-President George Bush stood on the deck of the USS Abraham Lincoln and declared the end of combat operations in Iraq following Operation Iraqi Freedom. The Internet Archive captured the subsequent press release on 6 May 2003.[3]

When it became clear that combat would continue in Iraq, the press release was modified on WhiteHouse.gov. A capture from 1 October 2003 shows the change.[4] Despite apparent attempts to modify the narrative of historical events, the archived web provides a critical record of actual history; the headlines in the two screen captures show the attempt to change the narrative, as the headline was modified to

read ‘major combat operations’, as opposed to ‘combat operations’ in the original press release.

The web is a living history; changes to the web are a story in and of themselves. As this chapter has illustrated, the web is also a tool for exploring the nature of change over time. Whether the focus is on a single story, a single newspaper or an entire ecosystem of information, there is validity and power in being able to trace history through the web.