ONLINE U.S. NEWSPAPERS AND THE PRESENCE OF BLOGOSPHERE CHARACTERISTICS

by
Anne M. LeBel

A Master's paper submitted to the faculty
of the School of Information and Library Science
of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Master of Science in
Information Science.

Chapel Hill, North Carolina

April 2005

Approved by

_____

Helen R. Tibbo

Anne M. LeBel. Online U.S. newspapers and the presence of blogosphere characteristics. A Master's Paper for the M.S. in I.S. degree. April, 2005. 48 pages. Advisor: Helen R. Tibbo

The purpose of this study is to research if online U.S. newspapers are employing key characteristics distinctively associated with the blogging community, and whether or not implementation can be predicted by a newspaper's market size or if the newspaper is a national paper.

A content analysis of 105 online U.S. daily newspaper websites found that online newspapers are not implementing a majority of blog-based features; they are however, implementing some of the most important − specifically free archives and permalink. The research also proved that there are statistical differences between circulation size, national status and the extent of implementation. The findings show that larger and national based newspapers implement a greater number of blogosphere characteristics.

Headings:

Content Analysis

Newspapers

News/Internet Resources

Web sites

Table of Contents

Introduction 3
Literature Review 9
Research Questions 12
Methodology 13
Specific Findings 16
Global Results 21
Conclusion 23
References 24
Appendix I: List of Newspapers Reviewed 29
Appendix II: Coded Variables 32
Appendix III: Inter-coder Training Guidelines 33
Appendix IV: Inter-coder Reliability 34
Appendix V: Statistical Analysis 37

Introduction

As the Internet broadened its reach in the mid-90s, many experts discussed how the online medium would change the fourth estate into a two-way, super-highway of multimedia, interactivity and communication. It was generally recognized that online news would become an endless stream of video clips and "hypertext links to archival content and external data." (Burkowski, Shepherd & Watters, 1994, p. 5) Newspapers of tomorrow would "likely be a converged, seamless mix of text, audio and video that combine the advantages of the traditional newspaper's important power to time-shift information to suit readers' needs and desires and provide context for events..." (Kimber, 1997, p. 596) Online local newspapers were told to focus on using the unlimited space offered by the Internet to give stories more detail and depth. Recommendations included linking to more background information, providing outside links so stories could be connected to other information sources, enriching stories using audio and video and differentiating news items depending on the time of day. (Hofstetter, 1998) Supplementing stories with links to "narratives, maps, tables, figures, drawings and other data" was highly desired by the industry observers. (Breecher, 1998, p.4)

These interactive expectations were driven by critics who called on online newspapers to develop an entirely different product from their standard print edition. For a variety of reasons, primarily economics, many newspaper publishers had "opted simply to put the content of the 'paper' product online, only to discover that the online world has

its own, often mysterious ethos."(Lapham, 1995, p.2) Publishers were warned that they would not succeed if they simply re-created printed content online and were even provoked with the question of whether or not the future of the online newspaper would amount to "what we can conventionally think of as news."(Kimber, 1997) A study in 2002 concluded "US newspapers that publish electronic editions on the Internet do not appear to reinvent themselves online."(Barhurst, p.477) Web guru and author of the Alertbox email newsletter, Jakob Nielsen, stated that "online newspapers should be written solely for the web… [using] a writing style that serves the online medium." (Mostafa, 2000)

The most dramatic effect however, still surfacing today, had more to do with the idea that the audience was now an active participant. News online was no longer defined by the one-to-many traditional communication model. The role of the audience was now as both producer and consumer of information and as a result "a new many-to-many communication model has emerged. (Lapham, 1995, p.1)

Some argued that, "as people all over the world begin to produce and then share information…they naturally depend less and less on the information that flows from more traditional sources."(Lapham, 1995, p.2) Others conceded the model had changed, but suggested − as so happened in a 1995 round-table discussion regarding old and new media by industry experts − traditional media would remain an integral part of the news reporting process even with consumers as active audience members. "We're at the end of 100 years of the consumer being the victim of top-down communications, and the people don't want to be told what is news. They do need institutional memory, fact checking and point of view… But they want to participate in the process."(Ingle)

The Rise of the Blog

According to the second annual report by the Project for Excellence in Journalism, readership of blogs has increased 58% in just the last six months with about 32 million Americans turning to Web logs for information. (Rainey, 2005) “Less than 10 years ago, practically all media was still a one-way street. Then early self-published online journals, or Weblogs, began to let bloggers bypass the corporate media gatekeepers to say and show practically anything they could think of to tens of millions of computer users around the world.” (Rosenbloom, 2004, p.31) This past year bloggers broke several major news stories, most notably contributing to the apology by, and possibly even the early retirement of, Dan Rather who had done a report for CBS on false documents obtained regarding President Bush’s career in the National Guard.

A survey conducted by the Pew Internet & American Life Project found “what media watchers may already know (and perhaps fear): that the Internet and the rise of the blogger are expected to drive greater change in the news media and publishing industries than in any other sector in society.” (Zeller, 2005) Using survey results that were received from thousands of Internet experts, the Pew project concluded that most agree the industry will undergo “yet more dramatic change.” One expert from the survey wrote, “The most obvious effects on news media are the rise of Weblogs supplanting the public’s attention to traditional news media, and the slow death of newspapers due to erosion of mindshare by online influences such as news web sites, chat rooms, message boards and online gaming.” (Fox, Anderson & Rainie, 2005, p.3) Newspaper publishers are taking note and realize that while the main audiences “are [still] relatively small, it is likely that they will become main stream. (NAA, 2005)

The Apparent Parallax

Newspaper publishers and bloggers have similar passions about the news and information in general. It appears differences between the two arise from methods used to deliver that passion to the masses and sometimes even those lines can become blurred.

> "While the blogosphere dramatically increases individual control over content, it dramatically decreases the power and authority of the traditional general-interest intermediaries, notably newspapers, magazines and broadcasters. ...Weblogs are influencing traditional media, adding irreverent wit, skepticism, and a challenge to real-world attitudes and events. Even as Weblogs increasingly try to deliver information and opinions like traditional media… traditional media increasingly try to deliver like Weblogs. For the reader and blogger alike, it's still a matter of credibility, trust and reputation." (Rosenbloom, 2004, p.33)

An example of this convergence can be seen as U.S. newspapers begin to buy up Internet companies in hopes of merging online writing styles and traditional print. In fact, the New York Times bought About.com this past February − coincidentally a website that uses blogging technology − to bulk up their Internet operations. (Graybow, 2005) Other online newspaper sites are actually adding Weblogs to their online audiences. "The [Greensboro, N.C.] News & Record Web site features 11 staff written Web journals, or blogs, including one by the editor that answers readers' questions, addresses criticisms and discusses how the paper is run….The News & Record's blogs range from 'just-the-facts, ma'am,' to slightly spicy." (Simon, 2005)

Bloggers and online journalists share the same goal of personalizing content for readers; however, the standards for doing so can be quite different. Bloggers are, by natural definition, known for voicing personal views. For example, some are popular for their leftist ideas; others are routinely visited for their right-wing agenda. Quite the opposite is applied to traditional media, where the idea of objectivity has always been the most important element of the free press. As mentioned earlier, introducing a degree of

personalization − usually by localizing − has also been a popular idea for the online newspaper community. As early as 1995, publishers were told that the many-to-many news delivery model required them to interact with readers in a manner "way beyond traditional letters to the editor." (Lapham, 1995, p.3) Community discussion was heralded as a way of localizing content. At a 2003 National Newspaper Association conference, participants were urged to build communities by using "local voices." (Murray, 2003) Smaller newspapers especially, were told to keep it local by providing a strong and "vociferously out-spoken" voice in the community. (Bressers, 2003)

As bloggers have always embraced new technology and the fast-paced nature of the Internet, for traditional media sites, there has been a fear of immediacy the Internet brings to the world of reporting the news. Maury Beecher wrote in 1998 that "online newspapers have a built-in speed advantage over print editions, yet hastily reported cyberscoops can prove embarrassing when they are wrong." (p.2) These days, news organizations commonly reference blog entries while at the same time bloggers are readily admitting when they get a story wrong. In fact, it usually becomes the main post of the day. Hugo Drayton, Managing Director for the Telegraph's new media unit, agrees that blogging has been great at "liberating people… but when people want the real take, they turn to trusted newspaper brands."(McCarthy, 2003, p.3) This sentiment can be paired with the paradox that "American consumers confront an ever-broader river of news from myriad sources… [and] the standard for gathering and presenting the information tends to be 'faster, looser and cheaper' than in the past."(Rainey, 2005)

The most recent battle between the two worlds has been over whether or not online newspapers should charge for archived content. Although revenues from online newspaper archives bring in very low ("single digit") percentages, opponents of opening

online newspaper archives argue it would hurt other revenue streams (including large database services) and if people are willing to pay, there is no reason not to charge. (Glaser, 2005) Others claim that opening archives will create additional advertising revenues, but more importantly, will solidify a newspaper's footprint on the Internet. "Now a rising chorus of voices is calling for… free archives at newspaper sites so that search engine and blogger links will remain live, newspapers can retain their authority in Google and articles can remain part of the online conversation." (Glaser, 2005) Even President and CEO of the Associated Press, Tom Curley, at a 2004 Online News Association conference spoke to the audience about this dichotomy and challenged them on how to "free" content from the "expensive containers" created by the industry while at the same time holding onto intellectual property. (Colombo, 2004)

Related to the issue of free, open archives is the idea of a permanent URL, or permalink, for each news story. "Web permanence is, I'd argue, one of the main things that journalists can learn from the more successful bloggers." (Waldman, 2005) In fact, web permanence is by far one of the most important tools of a blogger. In a 1998 Alertbox email newsletter, the term "linkrot" was defined as links that used to work but are no longer available. Jakob Nielsen's email stated that linkrot "contributes to dissolving the very fabric of the Web... [and] there is a looming danger the Web will stop being an interconnected universal hypertext and turn into a set of info-islands." This particular concept played and important role in forming the support for permalink, which is tremendously esteemed by the blogging community.

Former San Jose Mercury News technology columnist Dan Gillmor wrote "the consequence of the industry's tendency to push its older journalism behind a wall, where the initial links disappear and people now have to pay for access, is a loss of identity and

diminution of authority."(2005) Gillmor argues that newspaper stories should remain online as a "vital part of the permanent record of a community." He believes that online newspaper stories, of which most can be found nowhere else, "reinforce people's sense that the newspaper is a medium in their lives."

Simon Waldman, director of digital publishing for The Guardian (UK), discusses the value of permanence in detail as a guest writer for the PressThink Weblog.

> "Without permanence you slip off the search engines. Without permanence, bold ideas like 'news as conversation' fall away, because you're shutting down the conversation before it has barely started. Without permanence, you might be on the web, but you're certainly no *part of it*.
>
> Here's another example. Think of all the millions of words written by news organizations around the world about Abu Ghraib during 2004. Now go to Google and search…you will only find a handful of traditional media outlets mentioned in the first few pages… This isn't just a quirk in Google's search algorithm; this is about traditional media ceding responsibility for providing definitive, permanent record of major events. All that reporting effort, all that insight and expertise, all those contracts: now completely invisible to the millions who decide to use Google as their first and final tool for researching."(2005)

It is apparent that both the online newspaper industry and the blogging community have some common ground; however, differing opinions on the role the Internet should play in delivering their product seems to take center stage more often than not. This creates a perceived division in the shared objective – to deliver news and information – and it is the purpose of this study to investigate whether or not this is warranted.

Literature Review

The majority of research reviewed for this paper focused on the interactive nature of online newspapers and all were constructed using content analysis as the primary

methodology. A discussion of the researchers' objectives as well as their associated findings ensues in chronological order.

Gubman & Greer (1997) examined 83 newspaper websites to see if "criticism directed at the industry for failing to adapt to new technology [was] well-founded." The result of their research concluded that much of the criticism was warranted. Online newspapers were found to be duplicating print efforts on the web while not taking advantage of the immediacy the Internet offered. Gubman & Greer also found that interaction with readers, by email, was typically available and that larger newspapers tended to have more features available online. Updating content more than once a day, providing discussion forums and free archives and tailoring national news to local interest was lacking throughout all circulation groups.

Kamerer & Bressers (1998) analyzed a group of newspapers that included daily, non-daily and specialty online newspapers for technical and content features over a six month period. Their results showed that technical sophistication and content levels increased over that specific period of time. Daily newspapers were also found to be the most sophisticated, followed by non-dailies and specialized papers.

Tankard (1998) performed a content analysis on 135 websites and surveyed more than 400 daily newspapers to find out how frequently sites were being updated, to what extent hyperlinks, multimedia and extra online-only material was being used and the relationship of use to standard predictor variables like circulation and age of site. Results showed that the majority of online newspapers were updating their sites once a day, did not provide hypertext links and provided images on their sites but not multimedia objects. As for interactivity, most sites were found to provide email addresses for staff members and half were found to respond to reader emails. Half of the reporting sites also indicated

that most of their online content was not staff written. Circulation size of the newspaper was found to be the most reliable predictor of features provided online by the newspapers.

Massey & Levy (1999) studied the interactivity of English web versions of 44 newspapers in 14 different Asian countries. The researchers looked into the complexity of choices offered by the websites and questioned whether or not these newspapers were using the Net's potential. They also examined the extent consumers were allowed to create and add content to the newspaper's website. Massey & Levy found that the newspapers offered complexity in news content only, rarely "linking out of a daily news story to such background content as a same-day or archived article…" (p. 143). Other findings included letters to the editor as the most frequent way of allowing users to add content and two-thirds of the sites did not indicate immediacy of updated information on the site.

Peng (1999) explored various content and service trends in the Web newspaper publishing industry by asking what kind of presentation styles were present and how print and online versions differed. After reviewing 80 Websites and surveying 247 newspapers, Peng found that less than half did not provide hyperlinks to other news sites and only some allow "stories to stay on longer [than a day] or indefinitely." (p.6)

Also in 1999, Schultz conducted a content analysis on 100 U.S. online newspapers to find out how many newspapers offered email addresses to editors and writers, chat rooms, polls and surveys and discussion forums. Based on these features Schultz was able to gauge a level of interactive-ness and compare the levels based on newspaper ownership status and circulation size of newspaper. The researcher found that most papers did not offer email addresses to the authors, virtually none offered chat

rooms, a quarter offered polls and surveys and one-third offered forums. Larger newspapers were found to score higher on the index of interactive options.

Zeng (2002) studied factors that influence the level of interactivity by examining 106 U.S. daily newspapers. Interactivity was measured in terms of availability of email, online forums, section links, hypertext and the availability of a search engine. Results concluded that market size and age of online newspaper correlated to the level of interactivity. National newspapers were also found to be more interactive.

Zaharopoulos (2003) used content analysis on daily U.S. newspapers to find out what types of content was present and whether or not there was a relationship with circulation and the amount of content. Zaharopoulos found that there is a varying amount of content available and that "newspapers are still slow to utilize the full technological potential of the Web." (p. 6)

Weiss (2004) examined levels of innovation on 20 online newspaper homepages. The goal was to identify techniques being used by online newspapers and establish where they are in adoption process of certain practices. The research concluded that circulation size affects an online newspaper's level of innovativeness.

Research Questions

The purpose of this research paper is to investigate how much of the online newspaper industry and blogosphere actually coexist together. Previous research has tended to look at the interactive nature of online newspapers – some of these same characteristics are also examined here; however, in a different context. This is the first study to examine online newspapers by comparing them to strategies employed by the blogging community.

RQ1: To what extent are online U.S. newspapers implementing the strategies and characteristics found in the blogosphere?

RQ2: Is the extent of this implementation defined by the market size of the publication?

RQ3: Is the extent of this implementation defined by whether or not the publication is from a local or national market?

Methodology

To answer the above research questions, a content analysis of 105 U.S. daily online newspapers was conducted. The research sample was randomly selected from a listing provided by NewsLink.org, a reference site also used in some of the noted research studies (see Appendix I).

The sample was then categorized into one of four categories based on circulation numbers obtained from the *2004 Editor & Publisher Yearbook* for dailies. Sample sizes for each of the representative groups may seem skewed at first glance; however, according to recent reports issued by the Newspaper Association of America website, the sample size obtained in this study is truly representative of the industry. The 2004 NAA fact page states that a total of 1,456 daily newspapers exist in the United States, of which 85% percent fall under the 50,000 and under category. Adding in the 50,001 to 100,000 category depicts 93% of the U.S. daily market.

The unit of analysis for the study was categorized by story, sections offered, and the entire website of the online newspaper. Each category was coded for the presence or absence of a particular feature (see Appendix II). Characteristics were chosen based on descriptions provided by several experts in the blogging community who have either

written about what constitutes a blog or had made recommendations on how to improve web sites using different blogging techniques.

Co-author of the book *We Blog: Publishing Online with Weblogs*, Meg Hourihan, wrote an article for the O'Reilly Network describing blogs and used the following distinguishable characteristics in her article: frequent updates; ability for readers to comment; hyperlinks that accompany commentary; use of time stamps; and, permalink.

In 2003 Sébastian Paquet posted an article, still heavily linked to for reference purposes on KnowledegBoard.com, defining a blog as a website exhibiting these features: content owned and maintained by a single person; presence of hyperlinks; frequently updated; free access and archives available; and, permalink.

Dave Winer, from Weblogs at Harvard Law, wrote an article in 2003 called "What makes a Weblog a Weblog?" In it he describes the following as some of the main identifiable characteristics of a blog: archives; archive calendar and archive page URLs; permalink; comments; categories; syndication (i.e. RSS); email notification; rankings; pictures; media objects; membership/registration; and, discussion groups.

Jay Rosen, in a 2004 article posted on the PressThink journal, discussed the decision of a North Carolina newspaper (the Greensboro News-Record) to "ride with the outlaws" – in this case the community bloggers. In his article, he offered the following advice to online newspapers on how to work towards this same goal by implementing the following blog-like features: allow readers to comment after important articles and make the writer check often to additional discussion; change from never linking to always linking; offer a free archive; have journalists create, and consistently maintain, their bio pages (on stable URLs that are frequently linked to) based on who they are, where

they've been, etc…; and, keep a running list (rankings) of the most important issues to the community moving them off the list when the issue fades.

Based on these expert discussions, a matrix was designed (see Table 1) to identify the coding variables used in the research as well as assign point values once coding had been completed. The "Subscription Required or Other" feature was added by the researcher to capture those instances where features were not observed because a subscription was required or the website could not be accessed.

Table 1: Point Value Matrix

FeatureID	Feature	Hourihan	Paquet	Winer	Rosen	Point
0	Subscription Required or Other					0
1	Registration Required			X		1
2	Email (Bulletin) News Service			X		1
3	RSS Syndication			X		1
4	Timestamp	X				1
5	Image(s)			X		1
6	Media Objects			X		1
7	Comments	X		X	X	3
8	Frequent Linking - internal	X	X		X	3
9	Frequent Linking - external	X	X		X	3
10	Rapid Response - breaking local news	X	X			2
11	Rapid Response - breaking world news	X	X			2
12	Discussion Group or Forum Available			X		1
13	Writer, Columnist or Author bio/information page				X	1
14	Categorization of stories			X		1
15	Archive Calendar Available			X		1
16	Archive Page URLs			X		1
17	Rankings Displayed			X	X	2
18	Blog/Personl Voice		X		X	2
19	Permalink	X	X	X		3
20	Free Archive		X	X	X	3
					Total Possible Points	**34**

Note: Points were assigned based on the number of times an expert discussed the feature.

Inter-coder Reliability

Coding was conducted by the researcher and 3 fellow graduate students. Websites were accessed between the dates of February 9 and March 15, 2005 for initial coding and then checked for permalink by the researcher 30 days later. Exactly 20% of the sample was cross-coded. Coders were given a training session as well as guidelines to follow when coding (see Appendix III). To strengthen validity, these sites were coded on the

same date by the researcher and coder. As noted by McMillan, for content analysis, "…the primary challenge is to make sure that coders are actually cross-coding identical data. If Web sites are checked at different times by different coders…false error could be introduced." (2000, p.93)

Overall agreement between coders was 93%; however, because this can be considered an inflated number (since it does not account for chance agreements), Cohen's kappa was used as the true reliability measurement. As with other reliability calculations, the closer kappa is to 1.0, the higher the accuracy of the collected data. A threshold of .70 was used to conclude if inter-coder reliability was satisfactory and the overall kappa for the study was .76 (see Appendix IV).

Statistical Analysis

The software application SPSS was used for all statistical analysis. Chi-square tests were conducted on each individual feature (see Appendix V). Using the total accumulated feature points for each online paper, one-way ANOVA tests and the mean number of points were computed and discussed in the overall findings section.

Specific Findings

Subscription Required/Other (Feature 0)

There were 7 newspapers that required a subscription to view stories. One allowed for a free, one-day pass to view content on the site, and 2 of the sites allowed for only a couple of stories to be accessed. Four newspaper websites were unavailable at the time the research was conducted. Neither size of circulation or national status was determined to be significant factors.

Registration Required (Feature 1)

Fifteen newspapers required user registration in order to view stories on their websites, 3 of these were national papers. Both size of circulation and national status were significant factors.

Email News Service (Feature 2)

Twenty-seven online newspapers offered an email service to deliver news content. All five national newspapers provided this service. Both size of circulation and national status were significant factors.

RSS Syndication (Feature 3)

It was found that 4 national, as well as, 10 other papers provide RSS on their websites. Both size of circulation and national status were significant factors.

Timestamp (Feature 4)

There were 33 newspapers that had timestamps associated with stories. Three of these newspapers had the exact same date and time associated with every story, meaning all stories were probably uploaded at the same time. Neither size of circulation or national status was determined to be significant factors.

Images (Feature 5)

Fifty-nine, including all national papers, used images with their stories. Circulation size was not, but national status was, a factor.

Media Objects (Feature 6)

A total of 10, including 4 national papers, used media objects with stories. This feature was found to be the most significant based on circulation size and the second most significant for national papers.

Comments (Feature 7)

Eleven papers offered the user the ability to send comments about the article from within a story. No national papers offered this service. Only one paper actually posted comments received at the bottom of the stories. Circulation size was a significant factor, national status was not.

Frequent Internal Linking (Feature 8)

Seven papers (3 national) provided at least 2 internal hyperlinks, as part of a story, to others articles and areas of interest on their website. Both size of circulation and national status were significant factors.

Frequent External Linking (Feature 9)

Two newspapers, one national and one with a circulation size of approximately 28,000, provided at least two external links to other information sources from within an article. Circulation size was not a factor; national status was significant.

Rapid Response – Local News (Feature 10)

Thirteen papers, 4 national, provided breaking local news. For national papers, local meant content from their home market (i.e. Los Angeles for the LA Times, Washington D.C. for the Washington Post, New York for NYTimes.com, Boston for the

Christian Science Monitor and the United States was considered local for USAToday.com). Circulation size and national status were both significant factors.

Rapid Response – World News (Feature 11)

Sixty-three newspapers from the sample provided breaking world news (almost always from AP wire feeds). This feature was significant based on circulation size but not national status.

Discussion Group or Forum (Feature 12)

Fifty websites provided discussion forums for users not hosted by a third party. Most were in the form of guestbooks. Circulation size was, national status was not, a significant factor.

Bio Pages (Feature 13)

One national and 8 other newspapers offered biographies for their writers. This was determined to be a significant factor for circulation size but not for national status.

Categorization (Feature 14)

Forty-two websites offered a unique categorization for older stories or additional sections not related to the standard print sections. Categorization of stories was a significant factor based on circulation size, not national status.

Archive Calendar (Feature 15) & Archival URLs (Feature 16)

Nineteen newspapers offered a calendar to archived material. Most used the textual format. Fourteen of these papers used archival URLs. Circulation size and national status were not significant for either of these variables.

Rankings (Feature 17)

Nine papers (including all 5 national) offered some type of content ranking on their websites. Usually it was most emailed stories or most viewed. One site had most important stories for 2004. This feature was significant based on circulation size but not by national status.

Blogs (Feature 18)

Twenty-four websites offered one or more blogs written by their staff members, two of these were national papers. This was the second most significant factor based on circulation size, but was not significant with national status.

Permalink (Feature 19)

The most surprising result from the study is that permalink is the most widely implemented feature among the sample. A whopping 64% of the newspapers tested had valid links (after 30 days) for three headlines recorded from the original date of evaluation − although it should be noted that a longer evaluation period might produce different results. Two of the websites required that users sign-in to view the story. Ironically, 29 of these websites did not offer free archives and although users could get to the article with the saved URL, they were also asked to pay if searched for in the archive. Permalink was the least significant factor based on circulation size and national status.

Free Archive (Feature 20)

Forty-four online papers offered free archives, with an average of 4.69 years worth of stories available for the user. No national papers offered this feature. This was a significant factor based on circulation size. This was the only feature found significant without any implementation from papers in the > 100,000 category.

Global Results

To answer the first research question, the extent to which online newspapers are implementing blogosphere characteristics; the majority of features are not being implemented by online newspapers. With a maximum amount of 34 points possible, based on the constructed point value matrix, the most received by any one newspaper was 21. Three features were in use by the majority of the sample: permalink, breaking world news and images. The percentage of use for each feature can be seen in Table 2.

Table 2: Percentage of Feature Use

FeatureID	Feature	%
19	Permalink	64%
11	Rapid Response - breaking world news	60%
5	Image(s)	56%
12	Discussion Group or Forum Available	47%
20	Free Archive	42%
14	Categorization of stories	40%
4	Timestamp	31%
2	Email (Bulletin) News Service	26%
18	Blog/Personal Voice	23%
15	Archive Calendar Available	18%
1	Registration Required	14%
16	Archive Page URLs	13%
3	RSS Syndication	13%
10	Rapid Response - breaking local news	12%
7	Comments	10%
0	Subscription Required or Other	10%
6	Media Objects	10%
17	Rankings Displayed	9%
13	Writer, Columnist or Author bio/information page	9%
8	Frequent Linking - internal	7%
9	Frequent Linking - external	2%

It was found that market size does play a part in the extent of implementation. The mean number of points was smallest for the "25,000 or less" category and largest for the "100,001 or more" category (see Table 3). The larger standard error suggests that variation in total points associated with blog characteristics is greater for larger circulating publications. A one-way ANOVA test showed there is statistical significant difference in the mean number of points between circulation groups. There is also an extreme statistical difference between national and local papers. The mean number of points for a local paper was 8.17 and 17 for a national ($F = 20.988$; $p < .000$).

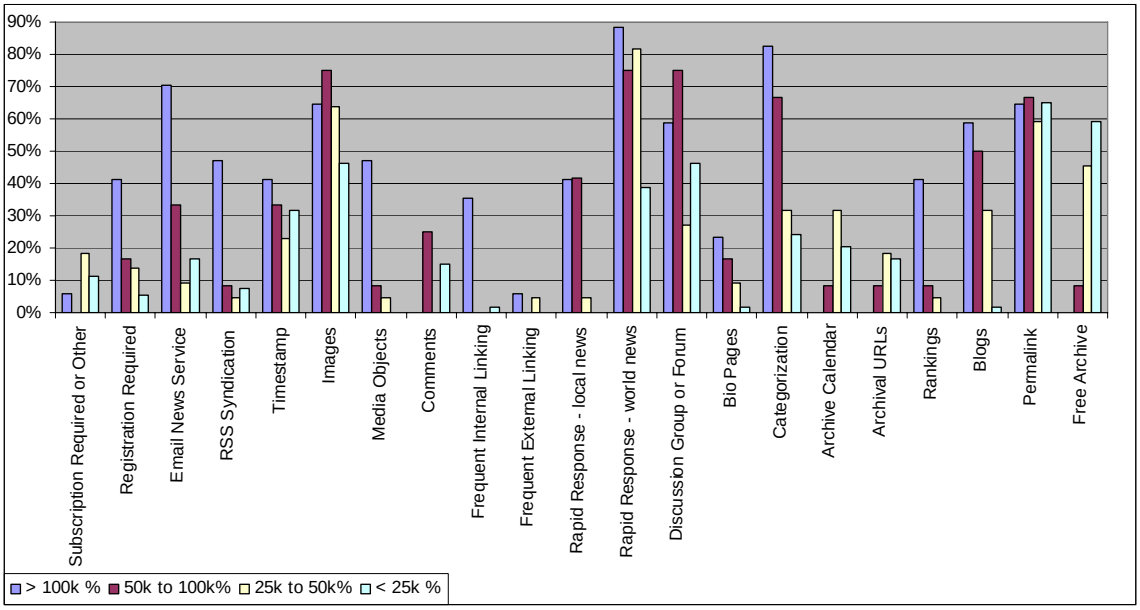
Table 3: Mean Points by Circulation Group

Circulation Group	Mean Points	N	Standard Error of Mean
25,000 or less	7.24	54	0.602
25,001 to 50,000	8.09	22	0.784
50,001 to 100,000	9.75	13	0.914
100,001 or more	12.71	16	1.14
Total	**8.59**	**105**	**0.448**

NOTES: F-ratio = 7.692, $p < .000$

Figure 1 shows the percent of features that were present for each circulation category.

Figure 1: Percentage of Present Features by Circulation Size



Conclusion

Needless to say, it is not surprising that larger and national based papers have more online blog features present; this has been proven time and again – using interactivity as the measurement factor – over the last decade as discussed in the literature review. In this case however, the defense that larger papers have more resources at hand is not a valid argument. Bloggers are, for the most part, individually-run operations. It can therefore be inferred that smaller circulating newspapers have just as much ability as larger-run publications to implement blogosphere characteristics.

In fact, as the research shows, the majority of smaller publications are implementing two of the most important features revered in the blogging community: permalink and free archives. This is extremely surprising since over the last year, online permanence, which is embodied by these two variables, has been the major issue blogging advocates use when criticizing newspaper websites – although, it is fair to say that bloggers would prefer it if the larger and national papers were the group that was embracing the idea of free archives.

Based on the researcher's findings, bloggers should be more concerned about the fact that online newspapers are not using the Web's most distinctive feature − hyperlinks − to enhance their online presence. Likewise, newspaper websites should take note that some of the more easily implement-able blog characteristics (specifically user comments with stories, writer biography pages and rankings) are absent from their Web publications. In conclusion, both bloggers and online newspaper publishers should realize that their two worlds coexist, and it is in everyone's interest that they work towards the common goal of solidifying user satisfaction when reading the news online.

References

Barnhurst, K. G. (2002, November 1). News geography & monopoly: the form of reports on US newspaper Internet sites. *Journalism Studies*, 3(4), 477-489.

Breecher, M. M. (1998). The newspaper of the future. *21stC*, 3.2. Retrieved January 26, 2005 from http://www.columbia.edu/cu/21stC/issue-3.2/breecher.html.

Bressers, B. (2003 May). Pinching pennies: Small papers find creative solutions in a slow economy. *The Quill*, 91(4), 20-23.

Burkowski, F.J., Shepherd, M.A. & Watters, C.R. (1994 October). Delivery of electronic news: a broadband application. Paper presented at the IBM Centre for Advanced Studies Conference, Toronto, Canada.

Colombo, S. (2004, November 12). Curley: 'The franchise is the content.' The Online News Association. Retrieved January 17, 2005 from http://journalist.org/2004conference/archives/000080.php.

Editor & Publisher. (2004). *Encyclopedia of the newspaper industry, 84th Edition*. Editor & Publisher: New York.

Fox, S., Anderson, J.Q. & Rainie, L. (2005, January). *The future of the Internet*. Pew Internet & American Life Project.

Gillmor, D. (2005, January 24). Newspapers: open your archives. *Dan Gillmor on grassroots journalism, Etc.* Retrieved March 19, 2005 from http://dangillmor.typepad.com/dan_gillmor_on_grassroots/.

Glaser, M. (2005, February 1). Pay or free? Newspaper archives not ready for open web…yet. *Online Journalism Review*. Retrieved March 19, 2005 from http://www.ojr.org/ojr/stories/050201/.

Graybow, M. (2005, February 18). U.S. newspaper chains pay up for web companies. *Reuters*. Retrieved February 18, 2005 from http://www.reuters.com.

Gubman, J. & Greer, J.D. (1997). An analysis of online sites produced by U.S. newspapers: are critics right? Apaper presented to the annual meeting of the Association for Education in Journalism and Mass Communication, Chicago, IL

Hofstetter, I. (1998). Multimedia applications for local newspapers and local information. *Computer Networks and ISDN Systems*, 30(1998), 1223-1232.

Hourihan, M. (2002, June 13). What We're Doing When We Blog. *O'Reilly Network*. Retrieved February 7, 2005 from http://www.oreillynet.com/pub/a/javascript/2002/06/13/megnut.html.

Ingle, B. (1995). Newspapers vs. online-versions. *Nieman Reports*, 49(2), p. 17.

Kamerer, D. & Bressers, B. (1998). Online newspapers: a trend study of news content and technical features. A paper presented to the annual meeting of the Association for Education in Journalism and Mass Communication, Baltimore, MD.

Kimber, S. (1997). The message is (still) the medium: the newspaper in the age of cyberspace. *Information Processing & Management*, 33(5), 595-597.

Lapham, C. (1995). Evolution of the newspaper of the future. *CMC Magazine*. Retrieved January 25, 2005 from http://www.ibilio.org/cmc/mag/1995/jul/lapham.html.

Massey, B. L. & Levy, M.R. (1999). Interactivity, online journalism, and English-language web newspapers in Asia. *Journalism and Mass Communication Quarterly*, 76(1), 138-151.

McCarthy, K. (2003, July 9). This is the future of online newspapers. *The Register*. Retrieved January 25, 2005 from http://www.theregister.co.uk/2003/07/09/this_is_the_future/.

McMillan, S.J. (2000). The microscope and the moving target: the challenge of applying content analysis to the World Wide Web. *Journalism & Mass Communication Quarterly*, 77(1), 80-98.

Mostafa, K. (2000, December 11). What's the use? A web guru wonders. *Editor & Publisher*, 133(50), p. 13.

Murray, G. (2003 September). Expanding the community newspaper through online news sites. Paper presented at the National Newspaper Association 117th Annual Convention, Kansas City, MO.

Nielson, Jakob. (1998, June 14). Fighting linkrot. *Useit.com alertbox*. Retrieved March 21, 2005 from http://www.useit.com/alertbox/980614.html.

Newspaper Association of America. (2004). Facts about newspapers. Retrieved March 18, 2004 from http://www.naa.org/info/factos04.

Newspaper Association of America. (2005 March). Media disintermediation: The rise of non-traditional media forms. *Horizon Watching*, 1-4.

Paquet, S. (2003 January 10). Personal knowledge publishing and its uses in research. *KnowledgeBoard*. Retrieved March 22, 2005 from http://www.knowledgeboard.com/cgi-bin/item.cgi?id=96934.

Peng, F.Y., Tham, N.I. & Xiaoming, H. (1999). Trends in online newspapers: a look at the U.S. Web. *Newspaper Research Journal*, 20(2), 52-63.

Rainey, J. (2005, March 14). Study warns of junk-news diet. *The Los Angeles Times*. Retrieved March 14, 2005 from http://www.latimes.com.

Rosenbloom, A. (2004). The Blogosphere. *Communications of the ACM*, 47(12), 31-33.

Rosen, J. (2004, December 18). More undercurrent: action in Greensboro on open source journalism. *Pressthink*. Retrieved January 17, 2005 from http://journalism.nyu.edu/pubzone/Weblogs/pressthink/2004/12/18/grns_nr.html.

Schultz, T. (1999 September). Interactive options in online journalism: a content analysis of 100 U.S. newspapers. Journal of Computer Mediated Communications, 5(1). Retrieved February 5, 2005 from http://www.ascusc.org/jcmc/vol5/issue1/schultz.html.

Simon, E. (2005, March 27). N.C. newspaper uses blogs to reach readers. *Associated Press*. Retrieved March 28, 2005 from http://news.yahoo.com.

Tankard, J. (1998). Online newspapers: living up to their potential? A paper presented to the annual meeting of the Association for Education in Journalism and Mass Communication, Baltimore, MD.

Tepper, M. (2003 September). The rise of social software. *netWorker*, 7(3), 18-23.

Treese, W. (2003 December). Why the newspaper has a great user interface. *netWorker*, 7(4), 13-14.

Waldman, S. (2005, January 7). The importance of being permanent. *Pressthink*. Retrieved January 17, 2005 from http://journalism.nyu.edu/pubzone/Weblogs/pressthink/2005/01/07/wldm_perm.html.

Weiss, A.S. (2004). The now what factor: the level of innovations among online newspapers. A paper presented to the annual meeting of the Association for Education in Journalism and Mass Communication, Toronto, Canada.

Winer, D. (2003, May 23). What makes a Weblog a Weblog? *Weblogs At Harvard Law*. Retrieved February 7, 2005 from http://blogs.law.harvard.edu/whatMakesAWeblogAWeblog.

Zaharopoulos, T. (2003). Online versions of U.S. daily newspapers: does size matter. A paper presented to the annual meeting of the Association for Education in Journalism and Mass Communication, Kansas City, MO.

Zeller, T. (2005, January 10). The Internet's Future? Depends on Whom You Ask. *The New York Times*. Retrieved January 10, 2005 from http://www.nytimes.com/2005/01/10/technology/10pew.html.

Zeng, Q. (2002 August). From print to online world: examining the predictors that influence the level of interactivity of newspaper's World Wide Web pages (Master's paper, Louisiana State University, 2002). Retrieved February 10, 2005 from http://etd02.lnx390.lsu.edu/docs/available/etd-0711102-130955/unrestricted/Thesis.pdf.

Appendix I: List of Newspapers Reviewed

A comprehensive list of 1,156 daily newspapers, organized by state, was obtained from http://newslink.org/daynews.html on January 18th, 2005. The data was placed in a Microsoft Excel worksheet, and six newspapers that are recognized as national papers by NewsLink were removed and the entire data set was randomly sorted. The top 100 rows were selected as the research sample and 5 of the 6 national papers were added back to the data set for a total research sample of 105 daily newspapers.

State	City	Newspaper	Circulation	Date Coded
Alabama	Anniston	Star	26,847	02/28/05
Arizona	Lake Havasu City	News-Herald	9,139	02/16/05
Arizona	Mesa	East Valley Tribune	95,041	03/08/05
Arkansas	Conway	Log Cabin Democrat	10,234	03/07/05
Arkansas	Fayetteville	Northwest Arkansas Times	18,017	02/28/05
Arkansas	Harrison	Times	9,424	02/18/05
Arkansas	Helena	World	2,580	03/10/05
California	Chico	Enterprise-Record	33,527	03/01/05
California	Davis	Enterprise	9,294	02/24/05
California	Fremont	Argus	31,823	02/21/05
California	Los Angeles	*Times	914,584	02/14/05
California	Red Bluff	News	7,140	03/09/05
California	Victorville	Press	29,989	03/07/05
California	West Covina	San Gabriel Valley Tribune	48,869	03/07/05
Colorado	Colorado Springs	Gazette	95,219	03/04/05
Colorado	Ft. Collins	Coloradoan	28,255	02/25/05
Delaware	New Castle	News Journal	116,398	02/21/05
DC	Washington	*Post	732,872	03/15/05
Florida	Daytona Beach	News-Journal	104,654	03/04/05
Florida	Palatka	News	11,739	03/07/05
Florida	Pensacola	News Journal	63,357	03/14/05
Florida	Stuart	News	99,224	03/10/05
Florida	West Palm Beach	Palm Beach Post	168,147	03/01/05
Hawaii	Kailua-Kona	West Hawaii Today	11,888	03/08/05
Idaho	Sandpoint	Bonner County Bee	4,537	03/08/05
Illinois	Centralia	Sentinel	14,313	02/09/05
Illinois	Danville	Commercial-News	15,525	02/18/05
Illinois	Quincy	Herald-Whig	23,719	02/23/05
Illinois	Rockford	Register-Star	66,249	02/16/05
Illinois	West Frankfort	American	3,510	02/23/05
Indiana	Hammond	Times	86,835	03/08/05
Indiana	Plymouth	Pilot News	6,227	02/28/05
Indiana	Terre Haute	Tribune-Star	27,499	03/09/05
Iowa	Ft. Madison	Democrat	5,057	03/02/05
Kansas	Coffeyville	Journal	4,103	03/02/05
Kansas	McPherson	Sentinel	4,227	03/02/05
Kentucky	Winchester	Sun	7,209	03/01/05
Louisiana	Franklin	Banner-Tribune	3,351	03/11/05
Louisiana	Lake Charles	American Press	35,704	03/09/05
Louisiana	New Orleans	Times-Picayune	253,610	02/09/05
Maine	Bangor	News	63,611	02/21/05
Maine	Portland	Press Herald	77,487	03/04/05
Massachusetts	Boston	*Christian Science Monitor	73,348	02/22/05
Massachusetts	Cape Cod	Times	51,378	03/08/05

State	City	Newspaper	Circulation	Date Accessed
Massachusetts	Lowell	Sun	48,571	03/04/05
Michigan	Holland	Sentinel	18,877	02/14/05
Michigan	Kalamazoo	Gazette	56,563	03/04/05
Michigan	Monroe	News	21,772	02/28/05
Michigan	Mount Clemens	Macomb Daily	44,699	02/09/05
Michigan	Royal Oak	Tribune	12,727	03/01/05
Michigan	Traverse City	Record-Eagle	28,574	03/08/05
Minnesota	Minneapolis	Pioneer Press	190,392	02/28/05
Minnesota	New Ulm	Journal	9,061	03/09/05
Minnesota	Rochester	Post-Bulletin	43,068	03/14/05
Mississippi	McComb	Enterprise-Journal	11,509	02/17/05
Missouri	Carthage	Press	3,543	02/15/05
Missouri	Kennett	Dunklin Democrat	4,985	03/01/05
Montana	Billings	Gazette	46,980	02/28/05
Nebraska	Beatrice	Sun	7,998	02/09/05
Nebraska	McCook	Gazette	6,339	03/08/05
Nebraska	Norfolk	News	17,282	03/09/05
Nebraska	York	News-Times	4,695	03/04/05
Nevada	Las Vegas	Review-Journal	160,391	02/21/05
New Jersey	Camden	South Jersey Courier-Post	77,031	03/01/05
New Mexico	Lovington	Leader	1,676	03/09/05
New York	Hornell	Tribune	6,472	02/16/05
New York	Ithaca	Journal	18,329	03/02/05
New York	New York	*Times	1,118,565	03/14/05
New York	Rochester	Democrat & Chronicle	173,900	02/23/05
North Carolina	Burlington	Times-News	27,368	03/08/05
North Carolina	New Bern	Sun Journal	15,391	02/21/05
North Dakota	Williston	Herald	4,919	03/10/05
Ohio	Cincinnati	Post	42,219	03/14/05
Ohio	Marietta	Times	11,241	02/15/05
Oklahoma	Duncan	Banner	7,929	02/22/05
Oklahoma	Oklahoma City	Oklahoman	207,538	03/10/05
Oregon	Bend	Bulletin	29,961	02/15/05
Oregon	Ontario	Argus Observer	7,049	03/01/05
Pennsylvania	Levittown	Bucks County Courier Times	65,243	03/09/05
Pennsylvania	Lock Haven	Express	9,274	03/02/05
Pennsylvania	Stroudsburg	Pocono Record	20,772	03/14/05
Pennsylvania	Wilkes	Citizen's Voice	43,244	03/14/05
South Dakota	Yankton	Press & Dakotan	7,887	02/14/05
Texas	Brenham	Banner Press	5,898	03/08/05
Texas	Galveston	Galveston County News	23,978	02/21/05
Texas	Houston	Chronicle	553,018	03/14/05
Texas	Huntsville	Item	5,646	03/11/05
Texas	Odessa	American	24,931	02/24/05
Texas	Pampa	News	5,240	03/03/05
Texas	Pecos	Enterprise	2,069	03/10/05
Texas	San Angelo	Standard-Times	26,966	03/09/05
Texas	Stephenville	Empire-Tribune	4,142	02/23/05
Texas	Texarkana	Gazette	31,262	02/14/05
Texas	Waco	Tribune-Herald	40,199	03/10/05
Texas	Wichita Falls	Times Record News	32,005	03/14/05
Utah	Salt Lake City	Salt Lake Tribune	134,985	03/03/05
Virginia	Manassas	Journal Messenger	7,109	03/14/05
Virginia	Martinsville	Bulletin	17,168	03/07/05
Virginia	McLean	*USA Today	2,154,539	03/08/05

State	City	Newspaper	Circulation	Date Accessed
Washington	Kent	South County Journal	46,038	03/08/05
Washington	Spokane	Spokesman-Review	100,587	03/02/05
Washington	Tacoma	News Tribune	128,511	03/04/05
West Virginia	Bluefield	Telegraph	19,137	02/21/05
West Virginia	Clarksburg	Exponent & Telegram	15,681	02/17/05
Wyoming	Sheridan	Press	6,565	03/04/05

*Defined as a national newspaper by NewsLink.

Summary of Distribution

By Geography:

The sample set was categorized into one of three possible geographic locations:

Northeast - Great Lakes	37 (35%)
South - Southeast	33 (32%)
West - Midwest	35 (33%)
TOTAL	105

By Circulation:

The sample set was categorized into one of four possible circulation types:

Less than 25,000	54 (51%)
25,001 to 50,000	22 (21%)
50,001 to 100,000	13 (12%)
Greater than 100,000	16 (15%)
TOTAL	105

Appendix II: Coded Variables

The following is a list of the variables used for coding each online newspaper. Variables are recognized as either present or not present (i.e. true/false). For the variable to be recognized as present, it must be maintained by the newspaper and on their website. Coders were instructed not to record variables of third party providers (e.g. calendars for obituaries on legacy.com, RSS for AP news wires, etc…). All Story variables must be found in staff written articles.

ID	Category	Variable
0	Website	Subscription Required or Other Issue – unable to view stories
1	Website	Registration Required
2	Website	Email (Bulletin) News Service – usually found if required to register or in the "About Us" section
3	Website	RSS Rendering or XML Available – usually noted at bottom of index page, in "About Us" or as it's own section
4	Story	Timestamp – posted with headline of story
5	Story	Image(s) – must be contained INSIDE (or with) story
6	Story	Media Objects – contained INSIDE story (i.e. animation, video, PDFs, etc)
7	Story	Comments – ability for the user to add or submit comments OR comments have already been posted by the author from readers about the story
8	Story	Frequent Linking – to internal sources (at least 2 per article - WITHIN AND regarding article only)
9	Story	Frequent Linking – to external sources (at least 2 per article - WITHIN AND regarding article only)
10	Section	Rapid Response – breaking local news *Note: There are three possible ways this can be identified: a) the website will state breaking (local) news; b) the timestamp registered is not from early morning (6-8am) but say mid-day; or c) an event occurs in real life (for example the Patriots win the Super Bowl) and the website reflects this reality immediately.*
11	Section	Rapid Response – breaking world news, usually AP news wires. *NOTE: Minimum of headline must be provided on newspaper's own website.*
12	Section	Discussion Group or Forum Available
13	Section	Writer, Columnist or Author bio/information page – not a listing of their stories but info page on them
14	Section	Categorization of stories other than standard sections of paper *Note: Usually by a subject or issue but NOT by date (that would be reflected by an archive calendar feature).*
15	Section	Calendar Available (in familiar format or textual) – used to locate past stories only, NOT events. *Note: Must be greater than 30 days and no manipulation required (i.e. search boxes).*
16	Section	Archive Page URLs – archived stories located through URL address Note: Example would be http://www.newspaper.com/archives/2003/07/05/. Usually this is found on websites that have a free archive and must be greater than 30 days.
17	Section	Rankings Displayed – of past stories or topic (Most Viewed, Popular, etc)
18	Section	Blog(s) – a personalized online voice from the paper
19	Story	Permalink *NOTE: Three headlines and their associated URLs were recorded during the coding process. Permalink was then tested 30 days later using the recorded URLs.*
20	Website	Free searchable or browse-able archive? *Note: A date feature was also recorded for the earliest available date in the archive.*

Appendix III: Inter-coder Training Guidelines

The purpose behind this project is to identify if key characteristics defined by the blogging community are currently being implemented on U.S. newspaper web sites and if the extent of implementation is affected by a newspaper's reach (national or local) and/or market share. As a coder you will have specific variables you need to identify. Variables must be provided directly from the newspaper and not from a third party source. To ensure accuracy between coders there are a couple of guidelines that must be followed:

1.) Websites are to be accessed after 10am Mon thru Fri. This is essential to verify if the publisher is posting breaking news for their community.
2.) A website must be reviewed for at least 15 minutes but no longer than 45 minutes. A minimum of 3 headlines and their URLs must be logged - unless the website is restricted by subscription or there are not at least 3 headlines published on the site. The 3 headlines should come from the front page (or local news section) of the newspaper and should be written by a staff member employed by the paper.
3.) Each coder has 7 websites that MUST be viewed on their scheduled date. It is important that this be done on the prescribed date since another coder will be reviewing the same website during the same time frame.

To begin your analysis, it is best to casually browse through the website to get a feel for general features. You might want to check the “About Us” section first to find out about the paper. Note to yourself if the website features XML or RSS capabilities or a news bulletin service.

Next would be to select three staff written articles. If a subscription or registration is required by the paper, it is usually when clicking on a headline users are instructed of this online policy. Assuming no registration is required, or it is and you have registered as a user, record the titles and associated URL of the story. Quickly skim the article and look for timestamps, images, media, user comments and hyperlinks (both internal and external).

Next, click through each of the section links and begin looking for other variables. Take note if breaking local or world news is offered. Is there a discussion or forum offered, do the writers and columnists have separate bio pages, has the paper put together special categorizations of stories (usually by story subject) or does the newspaper offer a list of stories by ranking (most popular or other)? If the newspaper has an archive, take a look at it to determine if the archive is free and searchable or browse-able, if there is a calendar of links to past stories available in familiar or textual format and if the archive is using archival URLs (see variable sheet for example). Verify the earliest date available in the archive by the provided calendar or statements made by the newspapers or by searching. Last, but not least, verify if the paper has a personalized online voice in the form of a blog or online journal maintained by the paper. You do not have to worry about the Permalink variable; this will be tested at a later date.

Appendix IV: Inter-coder Reliability

All coding used for statistical analysis was done by the author of this study. To test the validity of the research, twenty percent of the newspaper sample was randomly selected and coded a second time by three graduate students from the School of Information & Library Science at the University of North Carolina, Chapel Hill.

Coder 2 - Intercoding Reliability Test Data

Feature ID	Feature	Paper 1	Paper 2	Paper 3	Paper 4	Paper 5	Paper 6	Paper 7
0	Subscription Required/Other							
1	Registration Required	YesYes						
2	Email (Bulletin) News Servic	YesYes			YesYes		YesYes	
3	RSS Rendering or XML							
4	Timestamp		YesYes		NoYes			
5	Image(s)			YesYes		YesYes	YesYes	
6	Media Objects							
7	Comments	YesYes			YesNo		YesYes	
8	Frequent Linking (internal)							
9	Frequent Linking (external)			YesNo				
10	Rapid Response (local)							
11	Rapid Response (world)			YesYes		YesNo		
12	Discussion Group or Forum				YesNo	YesNo	YesNo	
13	Bio/information page					YesYes		
14	Categorization			YesNo			YesNo	
15	Calendar Available			YesNo				YesYes
16	Archive Page URLs							YesYes
17	Rankings Displayed							
18	Personalized Feature							
19	Permalink	N/A						
20	Free Archive						YesYes	YesYes

Variables Present Summary:									**TOTAL**
YesYes	**Coder 1=Y & Coder 2=Y**	**3**	**1**	**2**	**1**	**2**	**4**	**3**	**16**
YesNo	**Coder 1=Y & Coder 2=N**	**0**	**0**	**3**	**2**	**2**	**2**	**0**	**9**
NoYes	**Coder 1=N & Coder 2=Y**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**1**
NoNo	**Coder 1=N & Coder 2=N**	**17**	**19**	**15**	**16**	**16**	**14**	**17**	**114**
NOTE: Variables that both coders did not recognize are left blank.								**TOTAL**	**140**

Total % in agreement = Σa = (16 + 114)/140 = 130/140 = 92%

Cohen's Kappa

	Coder 1		
Coder 2	**Yes**	**No**	**row totals**
Yes	16	1	**17**
No	9	114	**123**
column totals	**25**	**115**	**140**

Sum of expected frequencies = Σef = (25*17)/140 + (115*123)/140 = 104.08

Kappa = (Σa - Σef)/(N – Σef) = (130 – 104.08)/(140 – 104.08) = 25.92/35.92 = .72

Kappa obtained is greater than the required .70. Reliability is satisfactory.

Coder 3 - Intercoding Reliability Test Data

Feature ID	Feature	Paper 1	Paper 2	Paper 3	Paper 4	Paper 5	Paper 6	Paper 7
0	Subscription Required/Other	YesYes					YesYes	
1	Registration Required							
2	Email (Bulletin) News Service							
3	RSS Rendering or XML							
4	Timestamp		YesYes		YesYes	YesYes		
5	Image(s)					YesYes		
6	Media Objects							
7	Comments							
8	Frequent Linking (internal)							
9	Frequent Linking (external)							
10	Rapid Response (local)			NoYes				
11	Rapid Response (world)			YesYes	YesYes			
12	Discussion Group or Forum					YesYes		YesNo
13	Bio/information page			YesYes				
14	Categorization					YesYes		
15	Calendar Available			YesYes				
16	Archive Page URLs							
17	Rankings Displayed							
18	Personalized Feature					NoYes		NoYes
19	Permalink				N/A			
20	Free Archive		YesYes				NoYes	

Variables Present Summary:

									TOTAL
YesYes	**Coder 1=Y & Coder 2=Y**	**1**	**2**	**3**	**2**	**4**	**1**	**0**	**13**
YesNo	**Coder 1=Y & Coder 2=N**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**1**
NoYes	**Coder 1=N & Coder 2=Y**	**0**	**0**	**1**	**0**	**1**	**1**	**1**	**4**
NoNo	**Coder 1=N & Coder 2=N**	**19**	**18**	**16**	**18**	**15**	**18**	**18**	**122**
NOTE: Variables that both coders did not recognize are left blank.								**TOTAL**	**140**

Total % in agreement = Σa = (13 + 122)/140 = 135/140 = 96%

Cohen's Kappa

	Coder 1		
Coder 3	**Yes**	**No**	**row totals**
Yes	13	4	**17**
No	1	122	**123**
column totals	**14**	**126**	**140**

Sum of expected frequencies = Σef = (14*17)/140 + (126*123)/140 = 112.4

Kappa = (Σa - Σef)/(N – Σef) = (135 – 112.4)/(140 – 112.4) = 22.6/27.6 = .82

Kappa obtained is greater than the required .70. Reliability is satisfactory.

Coder 4 - Intercoding Reliability Test Data

Feature ID	Feature	Paper 1	Paper 2	Paper 3	Paper 4	Paper 5	Paper 6	Paper 7
0	Subscription Required/Other						YesYes	
1	Registration Required		YesNo					
2	Email (Bulletin) News Servic	YesYes	YesNo	YesNo		YesYes		YesNo
3	RSS Rendering or XML							
4	Timestamp		YesYes		YesYes			
5	Image(s)	YesYes	YesYes	YesYes	YesYes			NoYes
6	Media Objects	YesYes						
7	Comments	YesYes						
8	Frequent Linking (internal)							
9	Frequent Linking (external)							
10	Rapid Response (local)	YesNo			NoYes			
11	Rapid Response (world)		YesYes	YesYes	YesYes	YesNo		YesYes
12	Discussion Group or Forum	YesYes		YesYes	YesYes	YesYes		
13	Bio/information page	YesNo						
14	Categorization	YesYes	YesYes	YesYes	YesNo			YesNo
15	Calendar Available	NoYes			YesYes			
16	Archive Page URLs				YesYes			
17	Rankings Displayed							
18	Personalized Feature	YesNo				YesYes		
19	Permalink				N/A			
20	Free Archive				YesYes			

Variables Present Summary:									**TOTAL**
YesYes	**Coder 1=Y & Coder 2=Y**	**6**	**4**	**4**	**7**	**3**	**1**	**1**	**26**
YesNo	**Coder 1=Y & Coder 2=N**	**3**	**2**	**1**	**1**	**1**	**0**	**2**	**10**
NoYes	**Coder 1=N & Coder 2=Y**	**1**	**0**	**0**	**1**	**0**	**0**	**1**	**3**
NoNo	**Coder 1=N & Coder 2=N**	**10**	**14**	**15**	**11**	**16**	**19**	**16**	**101**
NOTE: Variables that both coders did not recognize are left blank.								**TOTAL**	**140**

Total % in agreement = Σa = (26 + 101)/140 = 127/140 = 90%

Cohen's Kappa

	Coder 1		
Coder 4	**Yes**	**No**	**row totals**
Yes	26	3	**29**
No	10	101	**111**
column totals	**36**	**104**	**140**

Sum of expected frequencies = Σef = (36*29)/140 + (104*111)/140 = 89.92

Kappa = (Σa - Σef)/(N – Σef) = (127 – 89.92)/(140 – 89.92) = 37.08/50.08 = .74

Kappa obtained is greater than the required .70. Reliability is satisfactory.

Overall Agreement % and Kappa

Total % in agreement = Σa = (127+135+130)/(140*3) = 392/420 = 93%

	Coder 1		
Other Coders	**Yes**	**No**	**row totals**
Yes	55	8	**63**
No	20	337	**357**
column totals	**75**	**345**	**420**

Sum of expected frequencies = Σef = (75*63)/420 + (345*357)/420 = 304.5

Kappa = (Σa - Σef)/(N – Σef) = (392 – 304.5)/(420 – 304.5) = 87.5/115.5 = .76

Kappa obtained is greater than the required .70. Reliability is satisfactory.

Appendix V: Statistical Analysis

Subscription Required/Other (Feature 0)

Pearson Chi-Square = 3.203; p=.361

	By Circulation				
Value	< 25k	25k to 50k	50k to 100k	> 100k	**Total**
Not Present	48	18	12	16	**94**
Present	6	4	0	1	**11**
Total	**54**	**22**	**12**	**17**	**105**

Pearson Chi-Square = .614; p=.433

	By Status		
Value	Local	National	Total
Not Present	89	5	94
Present	11	0	11
Total	**100**	**5**	**105**

Registration Required (Feature 1)

Pearson Chi-Square = 13.463; p=.004

	By Circulation				
Value	< 25k	25k to 50k	50k to 100k	> 100k	**Total**
Not Present	51	19	10	10	**90**
Present	3	3	2	7	**15**
Total	**54**	**22**	**12**	**17**	**105**

Pearson Chi-Square = 8.96; p=.003

	By Status		
Value	Local	National	Total
Not Present	88	2	90
Present	12	3	15
Total	**100**	**5**	**105**

Email News Service (Feature 2)

Pearson Chi-Square = 23.782; p=.000

	By Circulation				
Value	< 25k	25k to 50k	50k to 100k	> 100k	**Total**
Not Present	45	20	8	5	**78**
Present	9	2	4	12	**27**
Total	**54**	**22**	**12**	**17**	**105**

Pearson Chi-Square = 15.167; p=.000

	By Status		
Value	Local	National	Total
Not Present	78	0	78
Present	22	5	27
Total	**100**	**5**	**105**

RSS Syndication (Feature 3)

Pearson Chi-Square = 20.104; p=.000

	By Circulation				
Value	< 25k	25k to 50k	50k to 100k	> 100k	**Total**
Not Present	50	21	11	9	**91**
Present	4	1	1	8	**14**
Total	**54**	**22**	**12**	**17**	**105**

Pearson Chi-Square = 20.192; p=.000

	By Status		
Value	Local	National	Total
Not Present	90	1	91
Present	10	4	14
Total	**100**	**5**	**105**

Timestamp (Feature 4)

Pearson Chi-Square = 1.543; p=.672

	By Circulation				
Value	< 25k	25k to 50k	50k to 100k	> 100k	**Total**
Not Present	37	17	8	10	**72**
Present	17	5	4	7	**33**
Total	**54**	**22**	**12**	**17**	**105**

Pearson Chi-Square = 1.989; p=.158

	By Status		
Value	Local	National	Total
Not Present	70	2	72
Present	30	3	33
Total	**100**	**5**	**105**

Images (Feature 5)

Pearson Chi-Square = 4.868; p=.182

	By Circulation				
Value	< 25k	25k to 50k	50k to 100k	> 100k	**Total**
Not Present	29	8	3	6	**46**
Present	25	14	9	11	**59**
Total	**54**	**22**	**12**	**17**	**105**

Pearson Chi-Square = 4.093; p=.043

	By Status		
Value	Local	National	Total
Not Present	46	0	46
Present	54	5	59
Total	**100**	**5**	**105**

Media Objects (Feature 6)

Pearson Chi-Square = 34.132; p=.000

	By Circulation				
Value	< 25k	25k to 50k	50k to 100k	> 100k	**Total**
Not Present	54	21	11	9	**95**
Present	0	1	1	8	**10**
Total	**54**	**22**	**12**	**17**	**105**

Pearson Chi-Square = 30.262; p=.000

	By Status		
Value	Local	National	Total
Not Present	94	1	95
Present	6	4	10
Total	**100**	**5**	**105**

Comments (Feature 7)

Pearson Chi-Square = 8.347; p=.039

	By Circulation				
Value	< 25k	25k to 50k	50k to 100k	> 100k	**Total**
Not Present	46	22	9	17	**94**
Present	8	0	3	0	**11**
Total	**54**	**22**	**12**	**17**	**105**

Pearson Chi-Square = .614; p=.433

	By Status		
Value	Local	National	Total
Not Present	89	5	94
Present	11	0	11
Total	**100**	**5**	**105**

Frequent Internal Linking (Feature 8)

Pearson Chi-Square = 26.831; p=.000

	By Circulation				
Value	< 25k	25k to 50k	50k to 100k	> 100k	**Total**
Not Present	53	22	12	11	**98**
Present	1	0	0	6	**7**
Total	**54**	**22**	**12**	**17**	**105**

Pearson Chi-Square = 24.0; p=.000

	By Status		
Value	Local	National	Total
Not Present	96	2	98
Present	4	3	7
Total	**100**	**5**	**105**

Frequent External Linking (Feature 9)

Pearson Chi-Square = 3.542; p=.315

	By Circulation				
Value	< 25k	25k to 50k	50k to 100k	> 100k	**Total**
Not Present	54	21	12	16	**103**
Present	0	1	0	1	**2**
Total	**54**	**22**	**12**	**17**	**105**

Pearson Chi-Square = 9.2; p=.002

	By Status		
Value	Local	National	Total
Not Present	99	4	103
Present	1	1	2
Total	**100**	**5**	**105**

Rapid Response – Local News (Feature 10)

Pearson Chi-Square = 31.357; p=.000

	By Circulation				
Value	< 25k	25k to 50k	50k to 100k	> 100k	**Total**
Not Present	54	21	7	10	**92**
Present	0	1	5	7	**13**
Total	**54**	**22**	**12**	**17**	**105**

Pearson Chi-Square = 22.12; p=.000

	By Status		
Value	Local	National	Total
Not Present	91	1	92
Present	9	4	13
Total	**100**	**5**	**105**

Rapid Response – World News (Feature 11)

Pearson Chi-Square = 21.163; p=.000

	By Circulation				
Value	< 25k	25k to 50k	50k to 100k	> 100k	**Total**
Not Present	33	4	3	2	**42**
Present	21	18	9	15	**63**
Total	**54**	**22**	**12**	**17**	**105**

Pearson Chi-Square = 3.5; p=.061

	By Status		
Value	Local	National	Total
Not Present	42	0	42
Present	58	5	63
Total	**100**	**5**	**105**

Discussion Group or Forum (Feature 12)

Pearson Chi-Square = 8.152; p=.043

	By Circulation				
Value	< 25k	25k to 50k	50k to 100k	> 100k	**Total**
Not Present	29	16	3	7	**55**
Present	25	6	9	10	**50**
Total	**54**	**22**	**12**	**17**	**105**

Pearson Chi-Square = .323; p=.570

	By Status		
Value	Local	National	Total
Not Present	53	2	55
Present	47	3	50
Total	**100**	**5**	**105**

Bio Pages (Feature 13)

Pearson Chi-Square = 8.976; p=.030

	By Circulation				
Value	< 25k	25k to 50k	50k to 100k	> 100k	**Total**
Not Present	53	20	10	13	**96**
Present	1	2	2	4	**9**
Total	**54**	**22**	**12**	**17**	**105**

Pearson Chi-Square = .875; p=.35

	By Status		
Value	Local	National	Total
Not Present	92	4	96
Present	8	1	9
Total	**100**	**5**	**105**

Categorization (Feature 14)

Pearson Chi-Square = 22.582; p=.000

	By Circulation				
Value	< 25k	25k to 50k	50k to 100k	> 100k	**Total**
Not Present	41	15	4	3	**63**
Present	13	7	8	14	**42**
Total	**54**	**22**	**12**	**17**	**105**

Pearson Chi-Square = .875; p=.35

	By Status		
Value	Local	National	Total
Not Present	62	1	63
Present	38	4	42
Total	**100**	**5**	**105**

Archive Calendar (Feature 15)

Pearson Chi-Square = 7.511; p=.057

	By Circulation				
Value	< 25k	25k to 50k	50k to 100k	> 100k	**Total**
Not Present	43	15	11	17	**86**
Present	11	7	1	0	**19**
Total	**54**	**22**	**12**	**17**	**105**

Pearson Chi-Square = 1.16; p=.281

	By Status		
Value	Local	National	Total
Not Present	81	5	86
Present	19	0	19
Total	**100**	**5**	**105**

Archival URLs (Feature 16)

Pearson Chi-Square = 3.842; p=.279

	By Circulation				
Value	< 25k	25k to 50k	50k to 100k	> 100k	**Total**
Not Present	45	18	11	17	**91**
Present	9	4	1	0	**14**
Total	**54**	**22**	**12**	**17**	**105**

Pearson Chi-Square = 1.16; p=.281

	By Status		
Value	Local	National	Total
Not Present	81	5	86
Present	19	0	19
Total	**100**	**5**	**105**

Rankings (Feature 17)

Pearson Chi-Square = 28.58; p=.000

	By Circulation				
Value	< 25k	25k to 50k	50k to 100k	> 100k	**Total**
Not Present	54	21	11	10	**96**
Present	0	1	1	7	**9**
Total	**54**	**22**	**12**	**17**	**105**

Pearson Chi-Square = 56.0; p=.000

	By Status		
Value	Local	National	Total
Not Present	96	0	96
Present	4	5	9
Total	**100**	**5**	**105**

Blogs (Feature 18)

Pearson Chi-Square = 32.0; p=.000

	By Circulation				
Value	< 25k	25k to 50k	50k to 100k	> 100k	**Total**
Not Present	53	15	6	7	**81**
Present	1	7	6	10	**24**
Total	**54**	**22**	**12**	**17**	**105**

Pearson Chi-Square = .875; p=.350

	By Status		
Value	Local	National	Total
Not Present	78	3	81
Present	22	2	24
Total	**100**	**5**	**105**

Permalink (Feature 19)

Pearson Chi-Square = .730 ; p=.866

	By Circulation				
Value	< 25k	25k to 50k	50k to 100k	> 100k	**Total**
Not Present	19	9	5	5	**38**
Present	35	13	7	12	**67**
Total	**54**	**22**	**12**	**17**	**105**

Pearson Chi-Square = .033 ; p=.856

	By Status		
Value	Local	National	Total
Not Present	36	2	38
Present	64	3	67
Total	**100**	**5**	**105**

Free Archive (Feature 20)

Pearson Chi-Square = 24.739; p=.000

	By Circulation				
Value	< 25k	25k to 50k	50k to 100k	> 100k	**Total**
Not Present	22	12	11	17	**62**
Present	32	10	1	0	**43**
Total	**54**	**22**	**12**	**17**	**105**

Pearson Chi-Square = 3.641; p=.056

	By Status		
Value	Local	National	Total
Not Present	57	5	62
Present	43	0	43
Total	**100**	**5**	**105**