

Seeking Sustainable Solutions to 21st Century News: A Case Study of Born-Digital Preservation

Jennifer L. Bonnet and Jennifer E. Moore

Introduction

Amid the rapid rise of born-digital news content, there is a pressing question on the minds of many librarians, archivists, researchers, and media workers: who is saving this digital draft of history? As institutions of social memory, with expertise in preserving and providing access to information, libraries and archives are struggling to find ways to capture proliferating born-digital news content, such as feature articles, blogs, infographics, and streaming video. This is an especially formidable task in an era of dispersed, and often ephemeral, digital-first publishing. Stemming from an industry in transition, this situation presents great challenges to cultural institutions that are attempting to ensure long-term preservation and access to digital documents that increasingly make up the historical record.¹ Moreover, little is known about the archival practices of news organizations regarding their digital-first content, raising additional concerns for future generations of scholars.

This paper presents an exploratory case study of issues related to born-digital news preservation among key stakeholders in the state of Maine: libraries and archival institutions, news organizations, and academic scholars.² This study seeks to better understand how each stakeholder group perceives the current state of born-digital news, as well as how they consider and/or attempt to address the problem of preserving born-digital news in the ever-expanding universe of digital information.

Literature Review

The twenty-first century has seen a dramatic shift in newsroom practice, from the production and circulation of newsprint (or a newsprint-digital combination) to largely digital-only publishing.³ With this transformation, the world is witnessing a stunning escalation of born digital news content, typically defined as news that originates in digital form rather than its analog predecessor. Such content poses a greater risk for potential loss than its print counterparts,⁴ with the National Digital Stewardship Alliance noting that, “the task of preserving born-digital web content has become progressively more difficult as websites become more complex and the scale of the web continues to grow.”⁵ This is especially true of sites that change rapidly, such as today’s online news, which is referred to by many names—digital first, web first, web centric, reverse publishing, native digital news, born-electronic news—⁶ and alarmingly, “at-risk news content.”⁷ According to *Taking Our Pulse: The OCLC Research Survey of Special Collections and Archives*, “[m]anagement of born-digital archival materials is still in its infancy” with cost and expertise emerging as major challenges to preserving this type of content.⁸ Reporting on the status of born-digital materials, Jackie Dooley of OCLC outlined five characteristics of concern for libraries and archives, namely that they are “undercollected,” “undercounted,” “undermanaged,” “unpreserved,” and “inaccessible.”⁹

Jennifer L. Bonnet, M.S.I., M.A. is Social Sciences & Humanities Librarian, The University of Maine, e-mail: jennifer.l.bonnet@maine.edu; Jennifer E. Moore, Ph.D. is Assistant Professor, Communication and Journalism, The University of Maine, e-mail: jennifer.e.moore@maine.edu

In this environment, stories of digital content loss are becoming more common. A community news site that relied on user-generated content was abruptly shuttered by NBC in 2013 before anyone could access the content for preservation.¹⁰ *The Columbia Missourian* lost 15 years of news stories and seven years of photographic material after a server crashed in 2002.¹¹ Making national news in 2014, BuzzFeed removed thousands of pre-2012 articles from its site “without any notice or disclosure” and with no record of what was withdrawn.¹² Even more recently, ransomware was used to attempt to extort a public radio station by threatening to destroy their archives if they did not supply a specified demand for digital currency.¹³ As Abbie Grotke, the Library of Congress Web Archiving Team Lead and Co-Chair of the National Digital Stewardship Alliance Content Working Group, warns, “With born-digital news content, if we don’t preserve it now before it goes away, it can and will be lost forever.”¹⁴ Implications for such losses will have long-term impacts on the quality and depth of the historical record.

Despite these challenges, there are numerous initiatives related to the archiving of born digital content that are serving as models for born-digital news preservation. The Library of Congress (LoC) has been archiving websites since the turn of the 21st century, and in recent years has expanded to include geospatial data, online forums, and blogs,¹⁵ not to mention social media platforms like Twitter.¹⁶ In conjunction with these larger efforts, local governments are working to develop effective methods for the preservation of their digital records.¹⁷ Additional support for these initiatives has emerged from the Internet Archive (IA), which has been preserving various aspects of the cultural and historical digital footprint for nearly two decades.¹⁸ The IA has been instrumental in archiving web, video, text, television, and audio content, as well as software. Like the LoC, they have begun partnering with local governments to answer the call for greater transparency around access to electronic information.¹⁹

Efforts among other institutions dedicated to the preservation of history have emerged in recent years,

with a focus on the preservation of emerging news content. In 2013, the National Digital Stewardship Alliance (NDSA) identified critical areas of concern specific to born digital news preservation.²⁰ The NDSA aims to work collaboratively to overcome barriers to preservation, whether financial, procedural, or technological. Their 2013 case study focused on building relationships among key stakeholders to preserve born digital news in hyperlocal and local contexts, i.e., areas in need of attention but often with fewer resources at the ready. In addition, digital news content has drawn increasing attention within national preservation efforts, leading to the design of best practices and guidelines for preparing such documents for long-term care,²¹ as well as case studies in born digital ingest workflows.²² While this work is an immense undertaking and solutions have yet to be identified that keep pace with the expanding journalistic record, initiatives like those emerging from the NDSA demonstrate the potential for partnerships aimed at capturing ephemeral news content before it is lost to history.

Building on the work of the NDSA and others, the current study examines the landscape of born digital news preservation in Maine, the authors’ state of residence. Specifically, this case study explores current practices and aspirations for preserving born digital news among key stakeholders: library and archival staff, news producers, and scholars. In 2014, the authors conducted the first phase of this multi-pronged study. Through a descriptive survey, library and archival staff, news producers, and researchers working in the state of Maine were asked about their perspectives and practices related to archiving news in the born-digital age.²³ Initial results of this study indicate that the long-term preservation of born-digital news is unevenly practiced and in need of a wide variety of resources to ensure success. These results will be published in the summer 2015 issue of *Newspaper Research Journal*.²⁴ The current paper will extend these findings by applying statistical tests to two key questions that examined 1) what stakeholders believe is important to archive and 2) what the institutional

groups (i.e., libraries, archives, news producers) need in order to carry out this type of preservation. This paper will also look at questions included in the survey that may prove useful to libraries and archives as they begin to examine their own user base regarding needs for long-term preservation, access, and use of born-digital news.

Methodology

Data Collection

Online surveys were designed and administered using Qualtrics survey software between May 19 and June 28, 2014. The authors distributed three separate surveys to library and archival staff, news staff, and university scholars in order to tailor questions to each group's unique roles. Each survey also asked a set of similar questions that were germane to all stakeholder groups. Library and archival staff were provided the same survey given the overlap in relevant questions and potential insights. For questions that respondents were unable to answer, there was a response option of "Not sure." In an attempt to minimize confusion or uncertainty about the meaning of the phrase "born digital," the following definition was included at the beginning of the survey: "For the purposes of this research, **born digital news** [emphasis original] refers to news content that originates in digital form." As an incentive, participants were invited to enter a raffle to receive one of four \$25 gift cards upon completion of the survey.

Participant Selection

Participants were preselected based on their roles as library and archival staff, news producers, and researchers working in the state of Maine. Recruitment strategies included emails to individuals in these stakeholder groups, as well as posts to social media platforms connected to the state's newspaper workers, researchers, librarians, and archivists. Academic librarian and archival listservs specific to the state of Maine posted the survey, and requests for participation were made at a Maine academic library conference. The Maine Press Association promoted the

survey throughout its organization, which included multiple appeals in its newsletter. Additional reminders were sent via e-mail. In an attempt to overcome single-digit participation among newspaper staff, personal e-mails were sent to editors and to the state newspaper association's board members to solicit their help in recruiting colleagues to participate.

In the spring and summer of 2015, the authors will conduct follow-up interviews with respondents who indicated a willingness to further participate in the study. These insights will strengthen our understanding of the pressing issues and needs faced by decision makers in our institutional groups (i.e., library and archival staff, news producers). Findings may also reveal potential partners for future born digital preservation efforts.

Research Questions

This research asks the following questions:

1. To what extent are library and archival staff, news producers, and scholars in the state of Maine attending to the preservation of born-digital news content?
2. What are the actions, if any, that institutions are taking to preserve born digital news content?
3. What challenges do institutional stakeholders face in terms of the types of born digital news content being preserved?

Results

The three separate surveys returned usable data from 113 library and archival staff, 34 newspaper employees, and 58 researchers. In the library/archival staff grouping, participants were primarily librarians from academic institutions, followed by a sizeable number of public library staff (see figure 1). In addition, respondents included archivists (with archival work as either their primary role or as a librarian/archivist combination), information technologists, and digital preservationists. The "Other" category included employees at school and state libraries, a historical society, and specialized libraries, as well as circulation, support, and managerial staff.

March 25–28, 2015, Portland, Oregon

FIGURE 1
Role Selection of Library/Archival Staff

Librarian	82
Archivist	4
Librarian/Archivist	7
Librarian/Digital preservationist	2
Librarian/Digital preservationist/Archivist	1
Information technology specialist	1
Librarian/IT specialist	4
Other	11
Librarian/Other	1

Among newspaper employees, most identified as editors or reporters, with additional respondents in the categories of photographer/multimedia producer, publisher, columnist, graphic designer, UX manager, or a combination of one or more of these roles. Researchers primarily comprised historians, although other disciplines included Anthropology, Psychology, Communication Studies, Political Science and Economics. A handful of additional fields were represented, including Native American Studies, Public Policy, Ecology, Sociology, and Literature.

RQ1. To what extent are library and archival staff, news producers, and scholars in the state of Maine attending to the preservation of born-digital news content?

Subquestion 1: compares the types of born digital news content each stakeholder considers important to preserve.

Subquestion 2: explores the regional born digital news coverage libraries/archives provide, news producers create, and scholars use in their research.

Subquestion 3: compares the types of born digital news content researchers use in their research with the types of born digital news content researchers consider important to preserve.

Subquestion 4: asks each stakeholder what organization(s) they believe should be responsible for preserving born digital news content.

In an effort to understand which types of digital news content were of utmost concern to each stakeholder, we asked about the importance of preserving a variety of news types based on a 5-item Likert scale (1 = unimportant, 3 = somewhat important, 5 = very important). All choices except one (digitized newspapers) referred to born digital content. “Digitized newspapers” was included in this question to further distinguish *born digital* news content from that which is more familiar (the digitization of analog newspapers), as well as to compare perspectives on the preservation of a format that may be more readily recognizable as an asset (i.e., digitized newspapers).

A one-way ANOVA comparing responses between stakeholder groups for each news content area revealed no statistically significant differences for the majority of choices: born digital news articles, news blogs, information graphics, tweets, Facebook posts, and online advertisements (see figure 2). Born digital news articles, news blogs, and information graphics were considered somewhat to very important across all stakeholders. Tweets, Facebook posts, and online

FIGURE 2
Subquestion 1: Importance of Preserving Born Digital News Content by Stakeholder Group (no significant differences)

Digital news content	Library/Archival Staff	News Staff	Researchers
Born digital news articles	4.1	4.1	3.9
News blogs	3.2	3.2	3.0
Information graphics on news sites	3.4	3.5	3.3
Tweets from news organizations	2.3	2.7	2.2
Facebook posts from news organizations	2.4	2.7	2.2
Advertisements from online news sources	2.1	2.5	1.9

advertisements were considered less important across all stakeholders.

However, there were significant differences among stakeholders at the .05 level in four areas: online polls, streaming videos, streaming video transcripts, and digitized newspapers. Of the four factors that were significant, a Levene’s test indicated that the assumption of homogeneity of variance had been violated for two of the four (video transcripts and digitized newspapers). Thus, we performed a Brown Forsythe test of equality of means to confirm significance, running all four comparisons to be certain. Given that three violated the Brown Forsythe test (streaming videos, video transcripts, and online polls), we ran Games-Howell posthoc tests, which are specifically designed to compare means for variables that do not meet the homogeneity of variances assumption.²⁵ Games-Howell revealed significant differences in responses to online polls, streaming videos, and streaming video transcripts, with digitized newspapers emerging as marginally significant at .063.

Specifically, library and archival staff considered polls and streaming video transcripts more important to preserve than did news staff, and library and archival staff found streaming videos and digitized newspapers more important to preserve than did researchers (see figures 3a and 3b).

To gain awareness of the regional focus of interest to scholars and news producers regarding born digital news, as well as the provision emphasis among library and archival staff, participants were asked about the born digital news coverage they use, create, or provide (see figure 4).

News staff most often indicated that they covered local and hyperlocal news, versus national, international, or other geographically-oriented news. Library and archival staff more often reported the provision of local news, followed by national and international news. Researchers tended to indicate a need for national news, followed by international and local news.

Five researchers in the “Other” category reported a need for state-level news, with additional “Other”

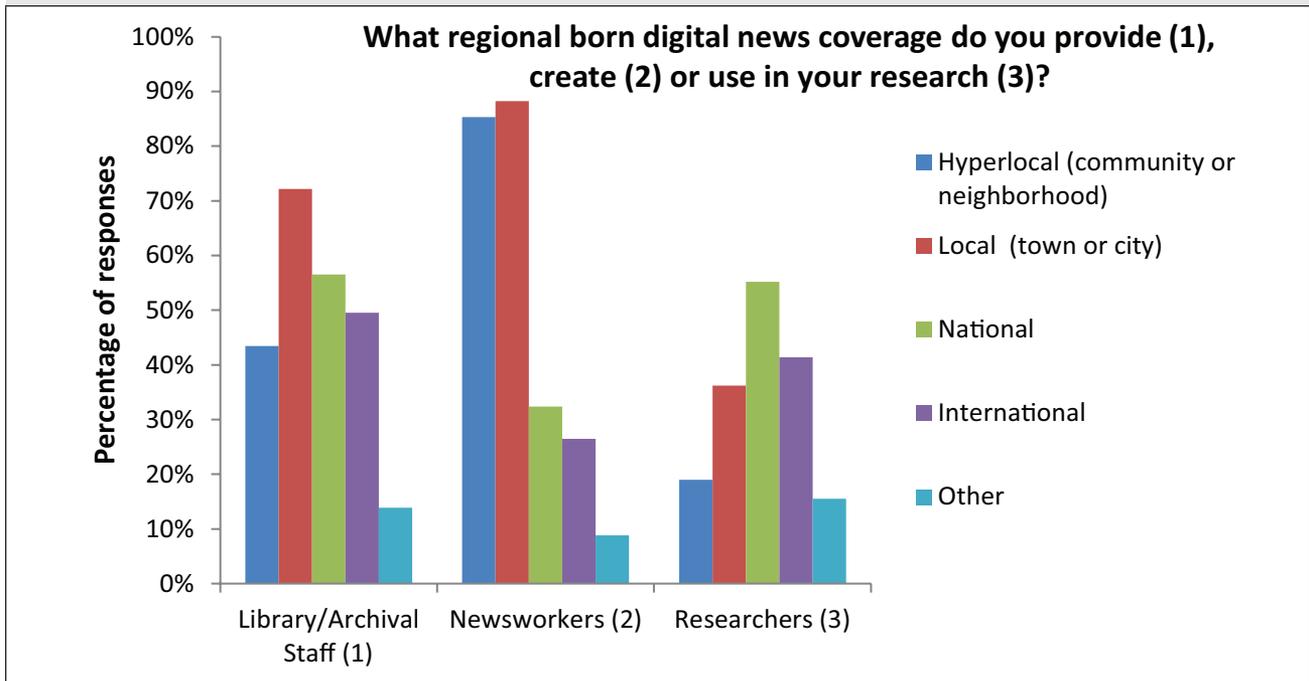
FIGURE 3A
Subquestion 1: Importance of Preserving Born Digital News Content by Stakeholder Group (Significant Differences)

Digital news content	Library/Archival Staff	News Staff	Researchers
Streaming videos	3.6	3.4	3.1
Streaming video transcripts	3.6	3	3.4
Online polls	3.1	2.5	2.9
Digitized newspapers	4.5	4.2	4.1

FIGURE 3B
Games Howell Post Hoc Tests (Significant Differences)

		Sig	95% Confidence Interval	
			Lower Bound	Upper Bound
Streaming Video	Library/Archival staff	.027	-1.01	-.05
	Researchers			
Transcripts of Streaming Video	Library/Archival staff	.013	.11	1.12
	News staff			
Online Polls	Library/Archival staff	.041	.02	1.30
	News staff			
Digitized Newspapers	Library/Archival staff	.063	-.93	.02
	Researchers			

FIGURE 4
Subquestion 2: What regional born digital news coverage do you provide, create, or use in your research?



mentions of New England, tribal/indigenous, and research news. Three newspaper employees also specified that they produce state news, and a handful of library and archival staff indicated that they provide access to state news. Additional responses from li-

brary and archival staff revealed that they did not know the regions their institutions covered, or only provided what patrons could access on the Internet (presumably the free web), with one mention of access through databases.

FIGURE 5
Subquestion 3: What types of born digital news content do researchers use in their research, compared with the types of born digital news content they consider important to preserve?

Born digital news content researchers use	Percentage	What researchers want preserved (1=unimportant, 3=somewhat important, 5=important)
Born digital articles	59%	3.9
News blogs	38%	3.0
Online polls	21%	2.9
Info graphics	28%	3.3
Tweets	3%	2.2
Facebook posts	5%	2.2
Streaming videos	14%	3.1
Video transcripts	12%	3.4
Online ads	0%	1.9
Don't use born digital content	29%	
Not sure	9%	

In a separate survey question, researchers were asked if they used born digital news specifically published in Maine for their work. Thirty replied in the affirmative (52%), ten said they did not, and one marked the “Other” option but did not provide an explanation.

Researchers were asked about the types of born digital news content they use in their work. Born digital articles were most prevalent, followed by news blogs, and information graphics. Twenty-nine percent indicated that they did not use any born digital news content in their work. When compared to what researchers indicated they wanted preserved, there were multiple matches, particularly at each end of the spectrum. Born digital articles received a high percentage of use and their preservation was deemed important, while social media posts were considered much less important and were used very little.

All stakeholders were asked to share their opinion regarding who should preserve born digital news. Figure 6 provides a comparison of the organization each stakeholder perceived should be responsible for preserving born digital news content.²⁶

Notably, all stakeholder groups most often selected news organizations as sites for preservation. News workers and researchers more often selected libraries than archives for the preservation of born digital news, whereas library and archival staff more often selected archives than libraries. “Other” responses among library and archival staff included third party publishers or vendors, the original creators of content, and the Library of Congress or National Archives. One respondent from the news worker group indicated uncertainty about who should be responsible, and another suggested the Internet Archive’s Wayback Machine. Among researchers, multiple respondents indicated that they simply did not know who should

preserve born digital news, while others suggested the Wayback Machine, database vendors, the free web, and non-governmental organizations.

RQ2. What are the actions, if any, that institutions are taking to preserve born-digital news content?

Subquestion 1: asks how long each institutional group has been archiving born digital news content.

Subquestion 2: asks each institutional group about their methods for storing born digital news content.

Subquestion 3: compares types of born digital news content scholars currently use in their research with born digital news content that institutions indicate they currently archive.

Respondents in institutional groups who indicated that they had knowledge of their organization’s preservation procedures were asked to identify when their organizations began preserving born digital news content. Figure 7 reveals that the library and archival staff who were able to answer this question (approximately 20% of the 113 respondents) more often responded that born digital news preservation had not yet occurred or efforts were less than three years old. Their news counterparts reported a longer window for born digital preservation efforts, with most respondents indicating that their organizations began preserving born digital news more than three years ago.

Both institutional groups were asked about their procedures for storing born digital content. Approximately half of the library and archival staff who answered this question were unsure of their institution’s procedures. However, several respondents mentioned the storage of files on servers, hard drives, through third-party services, and in the cloud. News staff

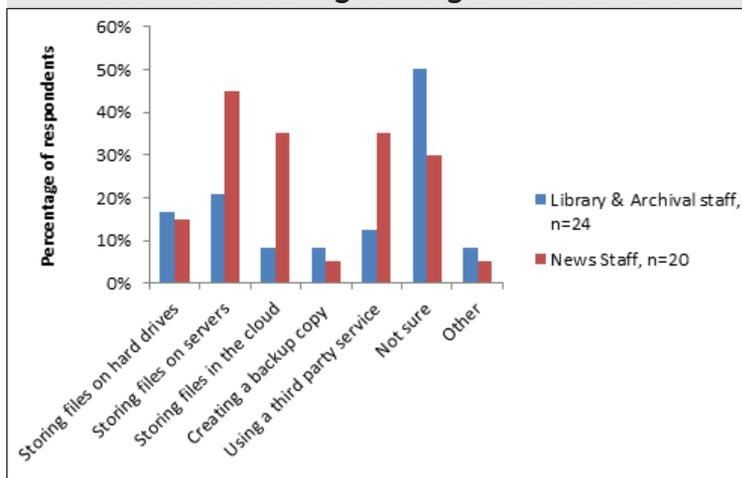
FIGURE 6
Subquestion 4: Who should preserve born digital news?

	Libraries	Archives	News Organizations	Other
Library & Archival Staff, n=113	49%	72%	82%	3%
News Workers, n=34	79%	62%	91%	6%
Researchers, n=58	66%	57%	69%	12%

FIGURE 7
Subquestion 1: Length of time organizations have been archiving born digital news content

	Library/Archival Staff, n=22	News Staff, n=20
Not yet	8	1
Less than 6 months ago	1	0
6 months to 1 year ago	2	2
1 to 3 years ago	7	3
More than 3 years ago	0	9
Not sure	4	5

FIGURE 8
Procedures for Storing Born Digital News Content



noted their predominance for storing files on servers, in the cloud, and through a third party service. Nearly one third of news staff were unsure of their organization’s procedures.²⁷

A larger number of news staff indicated that they preserved each news option listed, versus their library/archival staff counterparts. Over one third of library/archival staff indicated that they did not preserve any born digital news, and just under one third of researchers reported that they did not use born digital content in their research. The top five born digital news categories that received the most interest were, in order, articles, news blogs, infographics, online polls, and stream-

FIGURE 9
Subquestion 3: Comparison of born digital news content scholars currently use in their research with born digital news content institutions currently archive

Born digital content researchers use, n=58	Percentage	Born digital content archived ³⁴	Library/Archival Staff, n=23	News Staff, n=26
Articles	59%	Articles	22%	73%
News blogs	38%	News blogs	0%	50%
Online polls	21%	Online polls	0%	31%
Info graphics	28%	Info graphics	9%	50%
Tweets	3%	Tweets	0%	8%
Facebook posts	5%	Facebook posts	4%	8%
Streaming videos	14%	Streaming videos	9%	31%
Streaming video transcripts	12%	Streaming video transcripts	0%	0%
Online ads	0%	Online ads	9%	15%
Other	4%	Other	22%	4%
Don’t use born digital news	29%	Don’t archive born digital news	35%	0%

ing videos. These responses matched the rank order of content most news producers indicated they archive. Very little archiving was apparent by libraries and archives. Social media and online advertisements from news sites appeared the least important to researchers, which was matched by the organizations' current archival priorities.

RQ3. What challenges do institutional stakeholders face in terms of the types of born digital news content being preserved?

Subquestion 1: asks what resources organizations may need to preserve born digital news content.

Subquestion 2: asks for additional comments respondents would like to share about issues related to the preservation of born digital news content at their organizations, or more broadly

Participants were asked "In your opinion, what additional resources might your institution need to preserve born digital news content?" Responses were given a score of importance on a 5-point Likert scale (1=unimportant, 3=somewhat important, and 5=very important) and are displayed in figure 10. Space, equipment, standardization of format, and bandwidth were several of the "Other" responses. In addition, there were multiple mentions of preservation as mission creep.

An independent samples t-test revealed significant differences in four perceptions of need: financial support, interest in preservation efforts, education about best practices, and help identifying the type of content to preserve. Financial support and help identifying content to preserve were statistically significantly more important to library and archival staff, whereas an interest in preservation efforts and education about best practices were statistically significantly more important to news staff.

Subquestion 2: Please share additional comments about issues related to the preservation of born digital news content at your organization, or more broadly.

Further issues of concern emerged in responses to an open-ended question that asked participants to share additional comments that would help us understand issues related to the preservation of born digital news content at their organizations. Responses from the library and archival group appeared to emphasize resource constraints ("I know we should be doing SOMETHING, but no idea how to begin especially with zero resources to do so"), novelty ("There is so much that I would need to learn..."), collaboration ("Biggest roadblock (besides funding, of course) will be identifying and organizing digital efforts and responsibilities across multiple organizational types and locations"), and urgency ("Policies for preservation of born digital news are in their infancy. Loss of such content is a real risk.").

News staff shared comments that suggested a need for guidance in preservation efforts ("Best practices need to be developed and shared around the industry. Haphazard, organization-by-organization policies and procedures will do nothing for improving the archiving of digital news."), and funding ("We got serious about saving and promoting our archives when they realized we could possibly make money off it... If it helps offset the cost of doing it, then that might be the best selling point. Whatever it takes to preserve the history."), as well as an awareness of current preservation efforts for insular purposes ("While we preserve this, it's mostly for our own purposes and little thought it given to historical benefits or future needs.").

FIGURE 10
Subquestion 1: What resources might your organization need to preserve born digital news content?

	Library/ Archival Staff	News Staff
Financial support*	4.4	3.8
Qualified and trained personnel	4.2	4.2
Time	4.4	4.4
Interest in preservation efforts*	3.8	4.5
Education about best practices *	4.1	4.5
Help identifying the type of content to preserve*	4.1	3.6

* indicates significance at the .05 level

FIGURE 11
Subquestion 1: Independent Samples T-Test

Resources Institutional Groups Need for Preservation of Born Digital News		Levene's Test for Equality of Variances		t-test for Equality of Means			Confidence Interval
		F	Sig.	t	df	Sig. (2-tailed)	
Financial support	Equal variances assumed	4.353	.039	2.560	137	.012	
	Equal variances not assumed			2.259	41.564	.029*	.061 1.094
Interest	Equal variances assumed	3.727	.056	-2.605	135	.010*	-1.06 -.145
	Equal variances not assumed			-2.981	61.671	.004	
Best Practices	Equal variances assumed	4.143	.044	-2.057	137	.042	
	Equal variances not assumed			-2.565	78.289	.012*	-.705 -.089
What to Preserve	Equal variances assumed	.978	.324	2.120	137	.036*	.033 .936
	Equal variances not assumed			2.046	46.261	.046	

* indicates significance at the .05 level

Discussion

When asked about the importance of preserving certain types of born digital news, statistical testing revealed similarities among all stakeholder groups in several key areas. On average, all participants placed value on the preservation of more traditional born digital media like articles, and found some of the newer forms of content, such as information graphics and news blogs, at least somewhat important to preserve. News distributed via social media consistently ranked as less important for preservation purposes than the other options, possibly further reflecting an interest in preserving more traditional delivery mechanisms for the news. This finding could also point to the fact that social media platforms are often used to promote or drive traffic to a news organization's website and may not be prioritized as content. We anticipate that follow-up interviews will shed light on some of the preferences that emerged with this question. Statistically significant differences revealed that library and

archival staff placed more preservation importance on several news sources than other stakeholder groups. It is not surprising that staff who work in fields designed to safeguard cultural heritage materials and make them accessible to researchers would more often place importance on their longevity than others. Whereas library and archival staff indicated a near-statistically stronger preference than researchers for preserving digitized newspapers (i.e., scanned from their print counterparts), this category of digital news ranked highest among all groups, again pointing to a preference for a better known quantity and a resource with better known options for preservation than emerging forms of news. It is also possible that this finding illuminates the role librarians and archivists often play in attempting to anticipate future research needs through their collection and preservation activities.

Regarding the regional focus of news consumption, production, and access, local news predominated among the organizational groups in our sample.

This finding is not surprising in a state whose pride is reflected in its slogan, “Maine: The way life should be.” News producers’ coverage was highly localized, with over 80% of respondents pointing to the production of local or hyperlocal news coverage, and library and archival staff most often indicating the provision of local news for their library users. Increased attention to local and hyperlocal news mirrors a national trend,²⁸ and the importance of local news preservation is reflected in similar efforts to assess the future for born digital news.²⁹ Researchers in our sample reported a tendency to use national and international news in their research. However, local news ranked third among the choices provided and multiple scholars specifically mentioned their use of state news when offered the option of “Other.” It should be noted that 29% of researchers indicated they did not use any born digital content in their work. This finding may reflect our sample of researchers that was comprised primarily of historians. It is possible that these researchers have a tendency or preference for consulting digitized or print newspapers, rather than more recent, born-digital materials.

Scholars in Maine who indicated that they used born digital news were asked about the types of born digital content they consulted for their work and which content types they thought were key to preserve. Not surprisingly, more familiar content types such as articles and blogs were used by the largest number of respondents and ranked among the most important preservation options, with a smattering of use of other content. Online advertising and social media garnered the least attention from researchers and few indicated that they were important to preserve. Such insights may reflect current research agendas of this particular sample of researchers and may change over time as more of the cultural and historical record becomes comprised of born digital material.

Each stakeholder was asked which organizations they felt should be responsible for preserving born digital news. As noted in Moore and Bonnet,³⁰ the majority of respondents in all groups selected news organizations, which may point to expectations (or

aspirations) that those who are responsible for creating the content do or should also preserve it. Notably, news workers and researchers more often selected libraries than archives for preservation activities while the reverse was true for library and archival staff (i.e., they more often selected archives than libraries). These findings may reveal insider knowledge among library and archival staff of the distinct roles of libraries and archives, and possibly a greater familiarity for our sample of news staff and researchers with the work of libraries (versus archives) as information hubs. All stakeholder groups selected multiple organizations as keepers of born digital news, which may suggest an interest in duplicating efforts for preservation’s sake, or an uncertainty as to who would be the best option.

True to expectations among our stakeholders, we found that news staff preserved myriad types of born digital news, including articles, blogs, infographics, and more. It is unsurprising that as content producers, news organizations also hold onto much of their content. However, it is unclear if intentions for this content are for long-term preservation and/or access. This is of particular interest given the highly ranked need among news producers for best practices in preservation, as well as responses by news staff to the final, open-ended question that revealed constraints on time, finances, and awareness of how to move forward with preserving born digital content for a larger audience of users (i.e., not simply for internal purposes). As one respondent noted, “[Preservation is] definitely a back-burner issue that could use more attention, but with all the changes in the industry, our focus right now is on getting our product out and keeping it relevant.”

Whereas libraries and archives tend to be places researchers turn to access cultural and historical documentation, we found that either little born digital news content is being preserved or little is known about the practices among library and archival staff in our sample. This dearth of preservation awareness within libraries and archives could reflect a variety of factors, such as the incipient nature of born digital preservation efforts, the type of respondents who par-

anticipated in our surveys (there were many more self-identified librarians than archivists in our pool), or the environment in Maine itself around digital preservation preparedness. When asked about the length of time our institutional groups had been archiving born digital news content, results further revealed preservation nascence among libraries and archives in Maine. This finding was substantiated when only 20% of the sample of library and archival staff indicated that they were aware of their institution's preservation procedures, and of that number, 50% indicated that they were unsure of their organization's procedures for storing born digital content and 35% indicated that their organizations did not archive any born digital news.

Lastly, our institutional groups were asked what additional resources their organizations would need to preserve born digital news. Library and archival staff indicated that funding was a top priority, as was insight on what to preserve, the latter likely stemming from a service-oriented profession attuned to the needs of their library users. Knowing what users need consistently informs collection development, and in the case of born digital news, has implications for preservation decisions. Not surprisingly, news staff indicated needs related to best practices for preservation and an interest in the preservation of content. News workers are focused on producing the news, not necessarily safeguarding it. Thus, these findings open the possibility for cooperative engagement among both sets of decision makers to share their expertise in producing and preserving Maine news. While there were statistically significant differences among the needs of each organization, results revealed that all of the options averaged 3.5 or above across stakeholder groups. In other words, all survey options were ranked important to very important, revealing numerous needs yet to be met, and the potential for collaboration to meet these needs.

These findings have direct implications for the level of access to born digital news that will be available for future historians, cultural scholars, and journalists. This is an ongoing issue of concern for born

digital materials, given that "there are very few instances where users can access born-digital collections."³¹ Additional research is essential in order to identify opportunities for drawing on the greater community of cultural institutions and the media to look at the potential for shared solutions to the preservation of born digital news.

Limitations and Future Research

Survey Sample

Librarians and archivists responded to our survey at the highest rate, which is unsurprising given that this group has a vested professional interest in the preservation of cultural and social memory. Our lowest response rate came from news workers, who are likely more focused on producing the news than preserving it, with researchers in the middle. Responses by librarians outnumbered archivists by a large margin, which may have affected the paucity of respondents who indicated awareness of archival practices at their institutions.³² Further outreach to stakeholders in all categories in future iterations of this study will be essential to gain a more complete perspective on born digital news preservation in Maine.

Additionally, the focus of this study was limited to newspapers with online news operations. Opportunities for future research include an expansion of participation to include a greater range of news producers, such as television, radio, and other news outlets.

In the Center for Research Library's analysis of the preservation of electronic news, Reilly noted that "a viable approach to preserving electronic news will have to involve cooperation between libraries and newspaper producers, publishers, and/or aggregators."³³ Future research will need to account for vendors or aggregators as an additional stakeholder group, with the potential to incentivize born digital preservation through a funding model.

Follow-up Interviews

During the spring and summer of 2015, we will conduct follow-up interviews with respondents who indicated a willingness to further participate in this study.

We expect that these conversations will strengthen our understanding of the pressing issues and needs faced by decision makers in our institutional groups (i.e., library and archival staff and news producers). Additional outreach will take place among archival staff and news producers, two groups with the lowest sample sizes, to gain additional insights into the current state of born digital news preservation in Maine, and the potential for institutional partnerships going forward.

Notes

- Center for Research Libraries, "Preserving News in the Digital Environment: Mapping the Newspaper Industry in Transition," 2011, accessed February 15, 2014, http://www.crl.edu/sites/default/files/d6/attachments/pages/LCreport_final.pdf
- The survey data presented in this paper were originally collected for research that will be published in a forthcoming issue of the *Newspaper Research Journal*.
- Mark Jurkowitz, "The Growth in Digital Reporting," 2014, accessed December 15, 2014, <http://www.journalism.org/2014/03/26/the-growth-in-digital-reporting/>
- Ricky Erway, "Defining Born Digital," 2010, accessed February 13, 2014, <http://www.oclc.org/content/dam/research/activities/hiddencollections/borndigital.pdf?urlm=168879>
- National Digital Stewardship Alliance, "National Agenda for Digital Stewardship," 2014, accessed November 14, 2014, <http://www.digitalpreservation.gov/ndsadocuments/2014NationalAgenda.pdf>
- Ellen Sterling, "Digital First: What Does It Mean, And Where Will It Take Us?," 2011, accessed July 15, 2014, <http://www.editorandpublisher.com/Features/Article/Digital-First--What-Does-It-Mean--And-Where-Will-It-Take-Us-;> Paul Grabowitz, "Tutorial: The Transition To Digital Journalism," 2014, accessed December 10, 2014, <http://multimedia.journalism.berkeley.edu/tutorials/digital-transform;> Jurkowitz, "The Growth in Digital Reporting"; Bernard F. Reilly, Jr., "CRL preservation analysis of electronic news: Mapping the newspaper industry in transition," 2011, accessed February 1, 2015, <http://www.ingentaconnect.com/content/charleston/chadv/2011/00000013/00000001/art00017>
- National Digital Stewardship Alliance, "Case Study: Born-Digital Community and Hyperlocal News," 2013, accessed March 15, 2014, http://www.digitalpreservation.gov/ndsaworking_groups/documents/NDSA_CaseStudy_CommunityNews.pdf
- Jackie M. Dooley & Katherine Luce, "Taking Our Pulse: The OCLC Research Survey of Special Collections and Archives," 2010, accessed May 10, 2014, <http://www.oclc.org/content/dam/research/publications/library/2010/2010-11.pdf?urlm=162945>
- Jackie Dooley, "Demystifying Born Digital: "First Steps" Reports from OCLC Research," 2013, accessed December 10, 2014, <http://www.slideshare.net/oclc/demystifying-born-digital-first-steps-reports-from-oclc-research>
- Abbie Grotke, "Preserving Born-Digital Community and Hyperlocal News," 2013, accessed March 12, 2014, <http://blogs.loc.gov/digitalpreservation/2013/02/preserving-born-digital-community-and-hyperlocal-news>. NB: This site is up and running again.
- Edward McCain, "Content Matters: An Interview with Edward McCain of the Reynolds Journalism Institute," 2013, accessed November 20, 2014, <http://blogs.loc.gov/digitalpreservation/2013/09/content-matters-an-interview-with-edward-mccain-of-the-reynolds-journalism-institute;> Victoria McCargar Consulting, "Missouri J-School and the 'backstory,'" N.D., accessed November 20, 2014, <http://digitalpreservation.gov/documents/mccargar.pdf>
- Craig Silverman, "BuzzFeed's Ben Smith: 'We didn't fully think through' the removal of old posts," 2014, accessed August 30, 2014, <http://www.poynter.org/latest-news/regret-the-error/264007/buzzfeeds-ben-smith-we-didnt-fully-think-through-the-removal-of-old-posts>
- Joseph Lichterman, "How a virus demanding a bitcoin ransom almost destroyed a public radio station's archives," 2014, accessed November 30, 2014, <http://www.niemanlab.org/2014/11/how-a-virus-demanding-a-bitcoin-ransom-almost-destroyed-a-public-radio-stations-archives>
- Grotke, "Preserving Born-Digital Community and Hyperlocal News."
- Library of Congress, "NDSA Content Working Groups," accessed March 10, 2014, http://www.digitalpreservation.gov/ndsaworking_groups/content.html
- Matt Raymond, "How Tweet It Is!: Library Acquires Entire Twitter Archive," 2010, accessed March 15, 2014, <http://blogs.loc.gov/loc/2010/04/how-tweet-it-is-library-acquires-entire-twitter-archive;> Erin Allen, "Update on the Twitter Archive at the Library of Congress," 2013, accessed March 15, 2014, <http://blogs.loc.gov/loc/2013/01/update-on-the-twitter-archive-at-the-library-of-congress>
- Library of Congress, "Preserving State Government Information," accessed October 15, 2014, <http://www.digitalpreservation.gov/partners/states.html>
- "The Internet Archive is working to prevent the Internet—a new medium with major historical significance—and other "born-digital" materials from disappearing into the past." accessed October 10, 2014, , <https://archive.org/about>
- North Carolina Department of Cultural Resources, "State of North Carolina—Social Media Archive," 2014, accessed March 14, 2014, [http://nc.gov.archivesocial.com/;](http://nc.gov.archivesocial.com/) City of Austin, 2014, accessed March 20, 2014, [http://austintexas.gov.archivesocial.com/;](http://austintexas.gov.archivesocial.com/) Matt Enis, "South Carolina State Library Launches Social Media Library and Archive," 2013, accessed March 14, 2014, [http://www.thedigitalshift.com/2013/10/social-media/south-carolina-state-library-launches-social-media-library-and-archive/;](http://www.thedigitalshift.com/2013/10/social-media/south-carolina-state-library-launches-social-media-library-and-archive/) Maryland State Archives, "Maryland State Government Website Archive," 2014, accessed March 14, 2014, <http://msa.maryland.gov/msa/homepage/html/archive.html>
- National Digital Stewardship Alliance, "Case Study: Born-Digital Community and Hyperlocal News."
- Katherine Skinner & Matt Schultz, "Guidelines for Digital

- Newspaper Preservation Readiness,” 2014, accessed July 20, 2014, http://educopia.org/sites/educopia.org/files/publications/Guidelines_for_Digital_Newspaper_Preservation_Readiness_0.pdf
22. Digital Library of the Caribbean, “Born Digital Newspaper Preservation Workflows for the Florida Digital Newspaper Library and the Caribbean Newspaper Digital Library 2012,” 2012, accessed November 20, 2014, <http://dloc.com/AA00017121>
 23. A descriptive survey design aims to collect responses in order to explain current attitudes and conditions about a research topic. Roger D. Wimmer and Joseph R. Dominick, *Mass Media Research: An Introduction, Ninth Edition* (Boston, MA: Cengage Learning, 2009).
 24. Jennifer E. Moore and Jennifer L. Bonnet, “Preserving News as History: Perspectives on Born Digital News in the Pine Tree State,” *Newspaper Research Journal*, (2015, forthcoming).
 25. Laerd Statistics, “One-way ANOVA,” 2013, accessed January 5, 2015, <https://statistics.laerd.com/statistical-guides/one-way-anova-statistical-guide-4.php>
 26. Moore & Bonnet, “Preserving News as History.”
 27. Moore & Bonnet, “Preserving News as History.”
 28. For example, see Elizabeth K. Hansen and Gary E. Hansen, “Newspaper improves reader satisfaction by refocusing on local issues,” *Newspaper Research Journal* 32, no. 1 (2011): 98.
 29. The Reynolds Journalism Institute was recently awarded a \$35,000 grant from the Knight Foundation to develop a model for preserving born digital content at small newspapers. Jennifer Nelson, “Knight grant will help RJI develop born-digital-news preservation model,” 2014, accessed January 30, 2014, <http://www.rjionline.org/news/knight-grant-will-help-rji-develop-born-digital-news-preservation-model>
 30. Moore & Bonnet, “Preserving News as History.”
 31. Erway, “Defining Born Digital.”
 32. 72.5% identified as librarians and 10% as archivists with the latter solely marking the archivist option or a librarian/archivist combination.
 33. Reilly, “CRL Preservation Analysis of Electronic News.”
 34. Percentages reflect respondents who indicated that they knew their archival procedures.