

# FOCUS *on Global Resources*

Summer 2013 • Volume 32 • Number 4

## GLOBAL RESOURCES ROUNDTABLE: Access to News in the Digital Era

Mining Electronic News 2  
for Research

Broadcast and 5  
Multimedia News

Born Digital 8  
and Web-based News

A New Cooperative 11  
Agenda for News  
Preservation



### *In This Issue*

The current issue of *Focus* is a report on the recent [Global Resources Roundtable “Beyond the Fold: Access to News in the Digital Era.”](#)

Digital technology has profoundly altered the way news is produced and distributed, rendering traditional library strategies for ensuring the survival of this critical form of historical evidence obsolete. As mobile and web technologies become the primary channels for news distribution, and electronic databases supplant print and microfilm, the role that libraries play in supporting scholarly access to news is changing.

On Thursday, June 27, in conjunction with the ALA Annual Meeting in Chicago, CRL provided an opportunity for collection development and subject specialists from CRL libraries to explore current publishing practices and new scholarly uses of news content and to help formulate cooperative library strategies to support those uses. This report provides a capsule summary of the day’s discussions and presentations, and the outline of a new agenda for carrying CRL’s news preservation work forward.

The research library sector is uniquely positioned to play an impartial and informed role in ensuring the survival of critical documentation. Historically, CRL has taken a leading role in North America to preserve news for advanced research. Today, with the decline in government support for national libraries and funding agencies, it is incumbent upon the research libraries to step into the breach. Over the past three years CRL has concentrated major new resources on this challenge. The discussions and presentations at the roundtable produced a new blueprint for CRL’s activity in this realm in the years ahead.

—Bernard F. Reilly, Jr.  
*President*

Screen shot of Walter Cronkite on CBS,  
courtesy of Vanderbilt Television News  
Archive.

[www.crl.edu](http://www.crl.edu)

## Mining Electronic News for Research

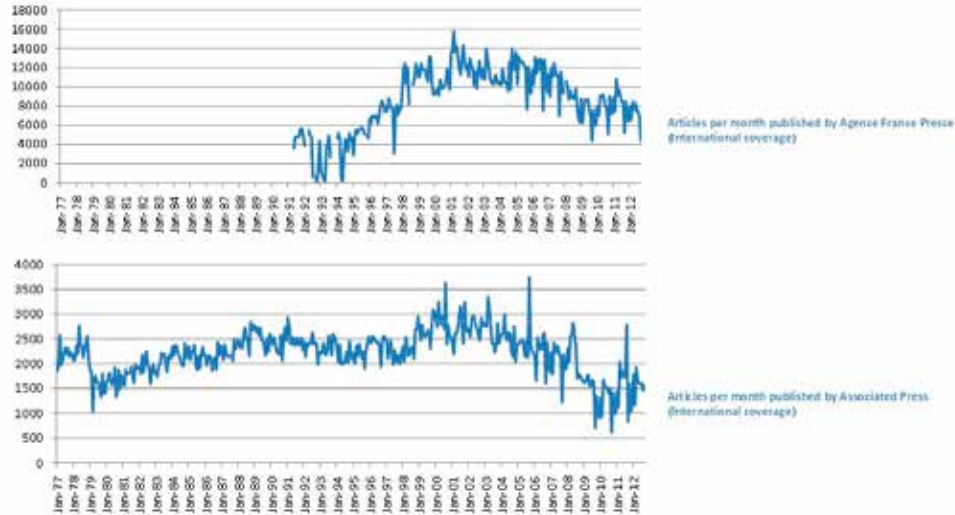


Chart from Kalev Leetaru's presentation shows articles published per month from various news sources.

**K**alev Leetaru, University Fellow, University of Illinois Graduate School of Library and Information Science, gave the keynote presentation, “Pioneers in Mining Electronic News for Research,” at CRL’s Global Resources Roundtable “Beyond the Fold: Access to News in the Digital Era.” Leetaru pointed out that technological developments of the past twenty years have revolutionized the means of communication throughout the world. (“Where there is power, there is Twitter”). The scale of media—especially social media—has grown to a nearly imponderable extent. New tools and systems have emerged to mine, store, and make available vast quantities of data for study. And data-computing methods, once only the domain of computer scientists, are now easily accessible to a wide range of researchers.

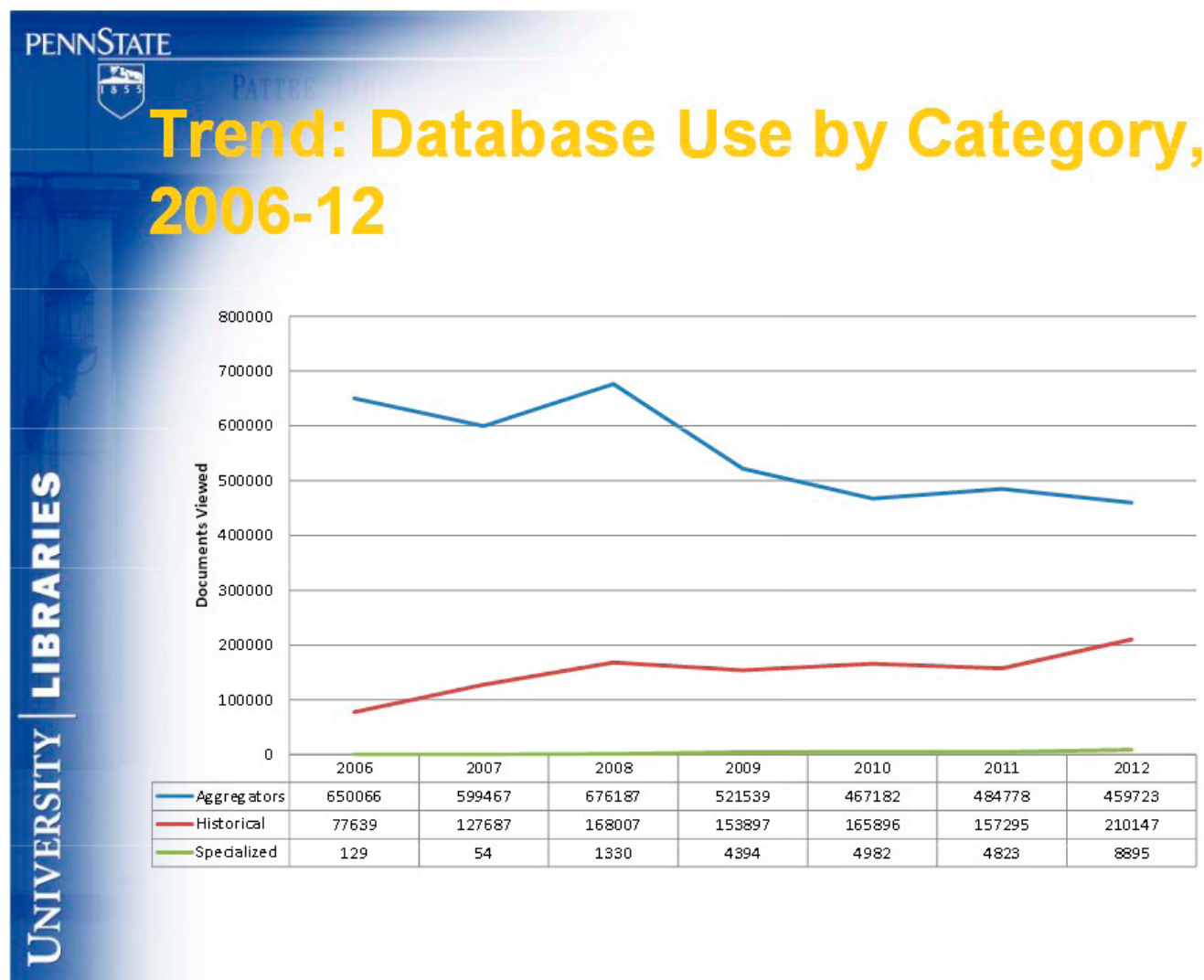
His major points:

- The volume of data—measured in number of articles available—from mainstream media sources such as *The New York Times* or *Agence France Presse* appears to be in a state of gradual and steady decline since the early 2000s. For researchers the explosion of web-based news presents immense opportunities as well as potential pitfalls. Researchers are drawing conclusions based on results from an exponentially increasing base of data. For instance while studies of violent demonstrations around the world suggest there is more instability now throughout the world compared to the mid-1990s, there may also be simply more widespread coverage of events than in decades past. Understanding the data in context—for example, the number of articles covering a particular event relative to the total number of sources and articles available—becomes important. Searching web-based content without a consistent and measurable baseline calls the research output into question.
- Traditional media has a relatively long time horizon, especially compared to the phenomenon of social media. Since the late 1800s, the study of the press and communications has been a perennial focus of historical research. Since the 1960s some new areas of academic research using news have come to the fore:
  - **Communications** focuses not only on what is reported, but also the context and medium of the reporting. Researchers in this area analyze the frequency of specific topics in reporting; reporting content, intensity, and valence; association of words and word patterns; and often perform advanced computational analysis.
  - **Political communication** is a specialized area of communications study, focusing specifically on biases, patterns, and trends in reporting on public figures, candidates, and political themes.

- **Political science/Sociology** research (which strongly emerged in the 1970s) utilizes coding schemes to categorize and assess the content of news coverage, creating data sets for ongoing research.
- **Linguistics** research often enlists computer scientists to provide and process large, controlled text corpora, where consistency is necessary to enable replication of findings. (Gigaword from the [Linguistic Data Consortium](#), for example, provides DVDs of news corpora over a set period of time for computational and linguistic analysis).
- Because of the different needs of research in the sciences, social sciences, and humanities, news must be preserved in a variety of formats, both as a “snapshot” of real presentation as well as a text-only set of information. All users, however, need to understand the “completeness” of the archive, e.g., what percentage of original articles from a given source are available. For web archives (born-digital news), we are still learning how researchers want to interact with archival content.

This slide from Debora Cheney's presentation depicts news database use (number of documents viewed) by aggregators, historical, and specialized.

**Debora Cheney**, Larry and Ellen Foster Communications Librarian for Pennsylvania State University (PSU) Libraries, provided a focused analysis of actual use of news databases. Her presentation, “[Use of Traditional News Databases at Penn State University: Trends and Implications](#),” was based on a study conducted at



Penn State University Libraries of student and faculty use of large news databases. Although the survey results are based on usage at PSU, they may suggest larger trends in the use of news at other academic institutions.

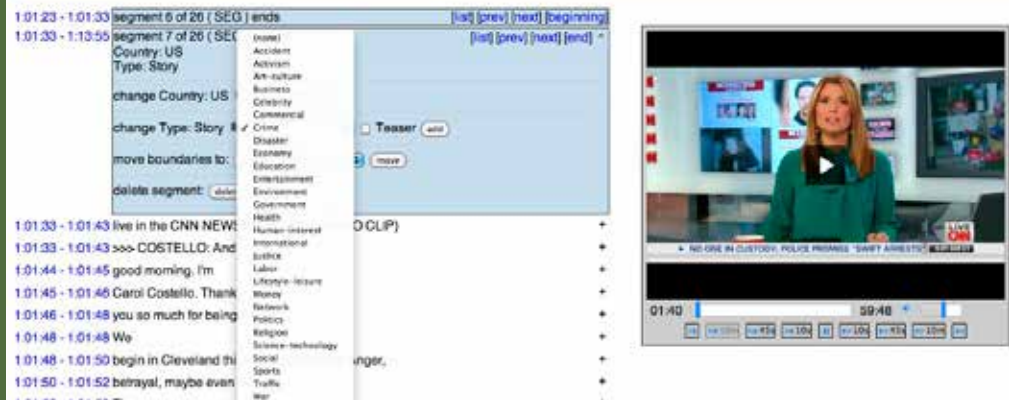
Cheney observed that, with the ubiquity of news on the web and the easy availability of the large databases, students don't necessarily associate libraries with news content anymore. She also noted that increasingly, researchers seek "news" from alternative sources: social media, blogs, and other content that the library does not supply.

Some key findings of the Penn State study:

- Overall use of the text-only aggregator databases at Penn State, such as LexisNexis and Factiva, is declining. Also declining is the viewing of articles from major newspapers like *The New York Times* and *Washington Post* within these products: The percentage of documents viewed from these key titles shows a steady decrease, from approximately 45% of all use in 2009 to just 13% in 2012.
- Despite the growing presence of international content in the aggregated databases, very few international sources are used to any great extent at Penn State. The exception is certain English-language resources, particularly financial publications, such as the *Financial Times*. Usage of wire service information (particularly business wires) and television news sources is increasing.
- Usage is generally much lower for historical news sources than for aggregated news, but use may be driven in different ways for historical sources. This use is affected by such factors as curriculum at Penn State and by patterns in deep historical research. Contrary to the aggregator trends described above, however, use of the historical databases seems to be increasing over time.
- New web-based library discovery systems seem to be having an impact on news database use. Resources not linked to discovery systems such as ProQuest's *Summon*, Ex Libris's *Primo*, and EBSCO *Discovery Service* must rely on direct knowledge and traffic to their product. Penn State noticed a major shift in use of *The New York Times* as delivered through *ProQuest Historical Newspapers* rather than *LexisNexis* or *Factiva* after implementing *Summon*. However, discovery systems bring with them added challenges for news and researchers:
  - Including news content with non-news content in discovery systems results in a greater number of search results, but a decline in the number of documents actually viewed.
  - Discovery systems do not distinguish "Page One" stories over other less-important content, nor do they necessarily weight results depending on the impact or the proximity of the source title (news is inherently a local phenomenon, but a system search may not distinguish between articles on the same subject from local or foreign sources).
  - The discovery systems also don't de-duplicate stories that appear on multiple news sources. Additionally, they don't search all resources in the same way. As a result, users may opt against using these tools for deep news research.

Cheney's analysis suggests that librarians can play an important role in helping researchers navigate the complex landscape of access to electronic news databases more effectively. Cheney noted that faculty members strongly influence student opinion on what resources to use, and that they should be encouraged to direct students to the resources provided by their own institution's libraries. ❖

## Broadcast and Multimedia News



This slide from the presentation by Sharon Farb and Todd Grappone shows searching options available in the *UCLA Library Broadcast NewsScape*.

The News Roundtable brought together representatives of two major television news archives to describe the services, resources, and benefits of their respective programs. As interest in television news broadcast content for research purposes grows, CRL is exploring how the community of academic libraries might better leverage, support, and exploit the community-based broadcast archiving efforts, namely the *Vanderbilt Television News Archive* and the recently announced *UCLA Library Broadcast NewsScape*.

As background, CRL presented [comparative data on the availability of broadcast news transcripts](#) in the major textual aggregator databases, and evaluated the potential usefulness of the Internet Archive's *TV Search & Borrow* for academic research.

The *Vanderbilt Television News Archive (VTNA)* at Vanderbilt University has been described as the “only comprehensive database of national television news.” As **Connie Vinita Dowell**, Dean of Libraries, and **Joseph D. Combs**, Associate Dean of Libraries, relayed [in their presentation](#), Vanderbilt began recording television news programs in 1968 when Paul Simpson, a Vanderbilt alumnus, discovered that the major U.S. networks were not keeping copies of their broadcasts. *VTNA* began taping the national network news (ABC, CBS, NBC), during that year's Republican National Convention. It added other programs and networks over time (CNN in 1995, MSNBC in 1996, and Fox News beginning in 2004) as well as special broadcast events (such as the Watergate hearings). *VTNA* now contains more than 40,000 hours of news and special broadcasts. Vanderbilt provides detailed metadata and abstracting of each story, with over one million records now available. Content is currently being captured in digital format, and nearly all of the legacy analog content has now been converted to digital, comprising approximately 200 terabytes of data and growing. Services include interlibrary loan of programs on DVDs and some materials are available for streaming over the web.

Access to the archive is constrained to some extent by copyright restrictions on re-broadcast or web distribution. Only NBC and CNN permit Vanderbilt to stream their content to subscribing institutions. Nonetheless the archive is used by researchers in a range of disciplines, including political science, sociology, history, and communications.

Research uses of the *VTNA* content include study of the media's treatment of particular events, trends, and subjects over time, and the impact of news broadcasts on public opinion or on “national conversations” about health care, climate change, elections, intelligence leaks, and so forth. To date the archive has largely been sustained financially by Vanderbilt University, with support from more than 130 sponsoring academic institutions.

[UCLA's Broadcast NewsScape is] driven by the paradigm shift in the ways that news is being produced as well as the way researchers use news.

As VTNA approaches its fiftieth anniversary, Vanderbilt is looking for ways to expand its support network: partnering with technology groups to develop video/audio “mining” capabilities and video and audio search tools. The university is open to developing partner models that allow “modest” cost recovery for the content providers, but the issues continue to be challenging. Working with CRL might enable broadening the base of support for the archives to encompass more academic libraries.

UCLA's **Sharon E. Farb**, Associate University Librarian, and **Todd Grapone**, Associate University Librarian for Digital Initiatives & Information Technology, presented on “*UCLA Broadcast NewsScape Archive of International Television News: A Transformative Approach to Using the News in Teaching, Research, and Publication.*” The **UCLA Library Broadcast NewsScape** is an international archive of over 200,000 news broadcasts recorded in analog and digital formats. Initially created by UCLA faculty members during the early 1970s, the archive was expanded in 2005, when professors in the university's Department of Communication Studies began capturing news programs in digital format. The content includes local news for Los Angeles, all national cable and broadcast news programs, and over a dozen international programs.

The project's methods are driven by the paradigm shift in the ways that news is being produced as well as the way researchers use news. The project team found that increasingly, scholars are interested in using media in advanced ways and linking to source materials, thus potentially raising issues of copyright.

UCLA is not, at present, able to stream content beyond campus. UCLA takes a different approach from Vanderbilt, hoping to exploit the exemptions to copyright restrictions afforded libraries by Section 107 of the U.S. copyright law, on the basis of the transformative nature of the UCLA platform for managing and disseminating the broadcasts. Through digital capture (including closed-captioning streams), metadata enhancements, and advanced tools for markup, searching, and playback (using facial and visual object recognition, story segmenting, tagging support), UCLA presents the collection in a way that is “transformative” or greatly enhanced for teaching and research.

UCLA aims to extend streaming video access to other UC system institutions. It might be feasible to eventually provide access for the broader academic library domain, in cooperation with CRL.

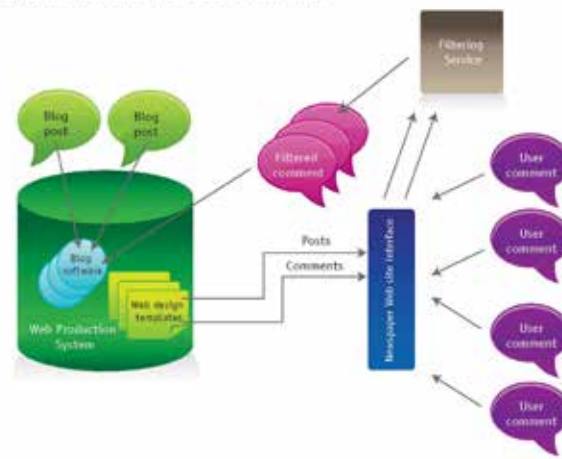
Inspired by Vanderbilt's model, but incorporating more advanced capturing techniques like *UCLA's NewsScape*, is the **Internet Archive TV News Search & Borrow**. Launched in 2012 (with content dating back to 2009), the archive is an open-access database of broadcast clips and text for search and discovery. As of June 2013, the *TV News Search & Borrow* contained a reported 452,000 broadcasts from 800 programs on 22 networks. Programs include national news broadcasts as well as local news programs (San Francisco, and later Washington, D.C.). The Internet Archive provides loans of programs on DVD and full viewing on the premises of the *Internet Archive* library.

Like *NewsScape*, the *Internet Archive* initiative captures closed-captioning streaming as the text archive, as opposed to official transcripts found in databases such as *LexisNexis*. Simple or advanced searching produces a textual snippet as well as a video clip that plays only brief (30 second) snippets of the segments, roughly aligned with the text results searched. No capture or download of text or image is permitted.

At present, the platform appears to work well with displaying the captured text and video segments from the basic search. However, the platform functionality lacks many features considered standard in academic databases. Live closed-captioning may contain numerous shortcuts, spelling errors, and textual ellipses. However, the wide accessibility of the content and the popularity of the *Internet Archive's* site make it likely that this tool will be useful for discovery purposes, leading potentially to additional uses of and interest in television news broadcast for research projects. ❖

## Born Digital and Web-based News

Lifecycle: Distribution (User-created content)



This diagram from Megan Bernal's presentation depicts a generalized distribution lifecycle of user-generated content for news organizations.

Having considered the challenges of print and broadcast media, the roundtable participants turned their attention to the issues surrounding born-digital news content. **Megan Bernal**, Associate Director for Library Information & Discovery Systems at DePaul University, spoke about the [production and distribution of electronic news](#) from the publishers' perspective, drawing on her background as Director of Information Services at the Miami Herald Media Company from 2005 to 2009.

Bernal provided an overview of the report issued in 2011 by CRL, "[Preserving News in the Digital Environment: Mapping the Newspaper Industry in Transition.](#)" That report outlined the "lifecycle" of news content published in newspapers and online, providing an overview of news workflow and production systems, and offered a basis for a rational and effective strategy for libraries to preserve news in electronic formats.

Among the findings in the detailed report, Bernal highlighted the complexity of the modern news organization, with many actors and systems contributing disparate streams of information to produce and distribute news content. The actors include the parent companies, which make deals with content providers and web-application providers to receive and distribute content, metadata, or other information at multiple levels. Interactive divisions are often separate units, usually business or advertising driven.

Traditional methods of acquiring and maintaining news are not sufficient to the task of capturing the electronic record. The diversity of sources (such as licensed and third-party content), dynamic nature of online content, and customized displays for individual users all work against capturing the "best edition" of an online publication. The proliferation of platforms, devices, and distribution networks used to read or access the news further complicates the ability to archive the user experience. Capturing a web-based PDF may capture the look and feel of the print edition, but these generally are not high-resolution copies and are stripped of much or all metadata generated in the production and layout systems and processes.

News enterprise systems employ a rich set of standards and metadata protocols for some types or levels of content that may be of interest to libraries and memory organizations. However, traditional news organizations are comprised of various departments (production, business, editorial, advertising, and circulation) that may produce metadata for their own particular purposes, distinct or siloed from the other parts of the organization.



There are numerous core enterprise systems within a news organization that implement separate production workflows and outputs (pagination systems and e-facsimiles; web production systems and web output, third party systems for additional content, etc.). Some systems may be more integrated than others. From an archival perspective, Bernal stated that the potentially highest impact “point of entry” for libraries would be within the editorial system, where the bulk of articles, metadata, (often) high-resolution photographs, and other related content is contained. She suggested that attempting to make a deal with news producers is a feasible, if not easy, approach to harnessing some of this content.

**Bernard Reilly**, President, Center for Research Libraries, presented on “[Legal Deposit Considerations in the Post-print Era](#)” In the print era, one or more copies of a work to be copyrighted by an author or publisher were often deposited in the national library. In return, the intellectual property of the work received certain legal protections. Since the 1940s, the Library of Congress and the British Library have accepted microforms as the primary form of legal deposit for newspapers to build their major collections of newspapers.

Since 2000, new legislation has been passed in a number of Western countries allowing for, and in some cases even requiring, legal deposit of electronic publications, including websites, in their respective national libraries. The legislation authorizes the deposit requirement to be fulfilled by one or both of two means: 1) deposit by the publishers; and 2) authorizing the national library to harvest directly from the web.

Reilly reported that despite these new statutory rights obtained by national libraries, actual legal deposit of electronic news is still limited. Many national e-deposit laws confined deposit to offline materials such as content on DVD and CD. Most libraries were not permitted to harvest website content that existed behind a pay wall, thus eliminating electronic news available from providers by subscription from capture. Other national libraries that targeted websites were harvesting periodic “snapshots” of the sites, sometimes as infrequently as once a year.

The British Library’s “Collecting Plans 2013–14,” call for harvesting “some 200 to 500 websites within scope . . . on a more frequent basis such as quarterly, monthly, weekly or even daily, in order to ensure that rapidly changing or updated content is archived adequately.”

Other libraries, such as the *Bibliothèque nationale de France*, are conducting e-deposit experiments with a major French regional newspaper, *Ouest France*, receiving comprehensive deposit of electronic news content directly. But this effort is limited to digital versions “exactly identical to the one distributed in printed form” and therefore does not capture the publisher’s web news output.

The U.S. Library of Congress (LC) does not yet have the statutory to require electronic deposit, nor the capability to ingest and archive news in electronic form on a regular basis. Some news sites are captured periodically as parts of occasional thematic web harvests done by LC or the Internet Archive. But these tend to be incomplete, crawled over periods of several days, and often lack critical content such as multimedia and database-driven content.

In instances where national libraries do harvest news content from the web, under current terms of deposit that content can only be made available within the confines of the depository library’s physical facility.

Reilly observed that the kinds of coverage and archiving of news content achieved in the print era through legal deposit programs are not likely to be provided in the

digital age. This suggests that the preservation of those materials for long-term use will have to be ensured by others.

**Mark Phillips**, Assistant Dean for Digital Libraries at the University of North Texas (UNT), described various initiatives underway at academic institutions including UNT aimed at the preservation of digital news content. The “Chronicles in Preservation” project is a collaboration of the [MetaArchive Cooperative](#), [Edu-copia Institute](#), [Chronopolis](#), and UNT, along with other academic institutions with collections of digital newspapers. The project aims to study and document the preservation practices of these institutions, and to model a distributed preservation framework to collaboratively preserve digital newspaper collections (both digitized and born-digital newspapers).

Participating institutions maintain a variety of digital news collections, including digitized historical newspapers, e-facsimiles of print newspapers, article “morgues,” and news websites dating back to the 1990s. These collections, predictably, contain an array of current and legacy content types and formats; are stored in a host of different systems; and have employed various metadata formats, OCR formats, and object identifier schemas over time. For many institutions, the size of the newspaper collections do not scale to the systems put into place for more traditional digital objects.

Based on surveys and in-depth interviews with participating institutions, the Chronicles in Preservation project is developing “[Guidelines for Digital Preservation Readiness](#)” to recommend specific practices to take advantage of available technologies and infrastructure that can bridge the gap in preservation readiness for institutions. The project team explored various means of “capacity building” and remediation at the partner institutions, which might be expanded to other institutions, publishers and/or content providers in subsequent phases of program activity.

Other deliverables of the Chronicles in Preservation project—still underway—include a comparative analysis of three leading distributed digital preservation approaches in the U.S., particularly in the context of preservation of newspaper holdings. The evaluation will include LOCKSS (implementation at the MetaArchive), iRODS (at Chronopolis), and CODA (at the University of North Texas).

The third primary deliverable will be the development of a set of “interoperability tools” to handle the exchange of content from partner repositories into the distributed preservation frameworks represented in this project. Details of the project are available through the project wiki at: [http://metaarchive.org/neh/index.php/Main\\_Page](http://metaarchive.org/neh/index.php/Main_Page).

The University of North Texas is engaged in a number of other efforts working to collect, digitize, preserve, and provide access to newspaper content. UNT coordinates the Texas Digital Newspaper Program, supported in part by the National Digital Newspaper Program (NDNP). In the past few years, UNT has been attempting to work with publishers to receive and archive current born-digital content for preservation. Numerous smaller newspapers, Phillips has found, are eager to work with memory organizations to “move their legacy forward.” By offering repository services for the current PDF facsimiles, UNT has found that the publishers often become more willing to collaborate on other efforts, such as the digitization of backfiles. UNT is presently working with the Texas Press Association and is offering services to other states’ historical archives. ❖

## A New Cooperative Agenda for News Preservation



Screen shot of a “CBS News Simulation” of the moon landing, July 21, 1969. Courtesy of the *Vanderbilt Television News Archive*.

At the end of the day, **Bernard Reilly**, President, Center for Research Libraries, returned to the key question of the roundtable, “What is the appropriate role for research libraries to play in ensuring the long-term accessibility and integrity of the journalistic record, in print, broadcast and the web?”

Reilly observed that a “tipping point” has been reached in libraries’ efforts to preserve news. Several things are contributing to this.

- In the realms of print and television news, space and other resources needed to maintain large bodies of paper and broadcast material are becoming scarcer, as is funding to continue to acquire large databases and research collections. Efforts by major research libraries that have historically supported the preservation of assets of national importance are approaching maturity and now require support beyond what their home institutions can provide.
- In the electronic era, the effective archiving of online news is simply not happening on a meaningful scale. The nascent legal deposit and web harvesting programs of various national libraries are either not yet scaled to archive significant amounts of electronic news content, or are not designed to capture online news content in formats that current research practices require. On the other hand, as Megan Bernal points out, impressive new capacity does exist in the media sector for managing and maintaining electronic news.
- At the same time, ever larger bodies of current and “historical” news content are becoming available in electronic format. The amount and variety of born-digital news have exploded in recent years and, as Debora Cheney’s survey revealed, navigating the myriad sources and platforms for access to current and non-current content presents daunting new challenges.
- As researchers’ capabilities for processing text and data grow, demand for access to these materials will only increase. Kalev Leetaru’s overview of text mining and other innovative uses suggests that researchers seek not just larger data sets but tools and functionality well beyond what libraries can themselves provide.
- And finally, research interest in television broadcast and cable news content, both transcripts and video, is growing, yet coverage of these in the current commercial offerings, particularly of international broadcast content, remains sparse.

Libraries will need to work strategically and at scale to accomplish the major tasks facing them. Preserving news is important not only for scholarship, but also for society. Reilly suggested the outlines of a cooperative agenda that research libraries might adopt to better ensure that the journalistic record will continue to be available and usable in the future. He said that CRL will pursue two complementary strategies in the news arena:

**An Advocacy Strategy:** CRL will build into its program some concrete measures to address the challenges identified in the roundtable:

- **Work to exert the collective influence of research libraries on news publishers and aggregators.** Libraries make up a sizable sector of the market for electronic news – by no means dominant but still sizable. As such, acting together they can probably do more to further the interests of scholars with the organizations that manage the content, whether large commercial database producers or media organizations.  
**CRL Action:** *Using this leverage in negotiations for member purchase of and subscription to electronic databases, CRL will work with NERL, JISC Collections, the Canadian Research Knowledge Network, and other appropriate partners to obtain greater functionality, exposure of more data on the contents and usage of news databases, greater interoperability among news platforms, and other concessions from the publishers, aggregators, and vendors of electronic news.*
- **Determine how the CRL community might better support the “non-proprietary platforms” that archive and provide electronic access to traditional news content.** An opportunity now exists to leverage capabilities and considerable historic investment by institutions like Vanderbilt, UCLA, and University of North Texas to create greater accessibility and functionality by broadening their base of support.  
**CRL Action:** *In FY 2014 CRL will increase its own investment in the digitization of foreign newspapers, a longstanding focus of CRL collecting.*
- **Enlist support from communities outside the traditional humanities and social sciences, i.e., business, law, and public policy libraries and researchers, in the effort to preserve international and foreign news content.** The recent [Global Dimensions conference](#) noted the growth of, and investment in, international studies in the professional schools of U.S. and Canadian universities. This new interest will create demand for news content from other world regions, and thus incentives for professional school libraries to invest in preservation and electronic access.  
**CRL Action:** *Factor the needs of those libraries into CRL’s strategic dealings with publishers and aggregators, and enlist them in funding electronic access to foreign news collections.*

**A “Meta-preservation” Strategy:** With scarce resources, libraries must concentrate on measures that are most likely to produce real benefits for researchers. Informed investment in preservation will require analysis and information about the landscape of news production, management and distribution, and use. Therefore CRL will build into its operations ways in which to:

- **Increase our understanding of the needs of the researchers we serve.** Determine how today’s researchers use news content, the kinds of tools and analytical practices they bring to bear in that use, and the kinds of scholarly products and outputs they produce. We know little about which of the many versions of a given “story” or news report is the most useful for researchers, and how useful “snapshots” of new websites are. Determining what content is appropriate will determine the type of preservation approaches required. Scholarly practices are changing rapidly, and current preservation models may not be adequate to serve today’s or tomorrow’s research.  
**CRL Action:** *CRL will continue to analyze and report on the practices of researchers using news corpora, through its webinars and [Primary Source Awards](#) program.*
- **Increase the amount and quality of data available on digitized news back files and the contents of news aggregator databases.** Libraries and commercial publishers are investing heavily in digitization of back files of newspapers. With little information available about what has been digitized, and what is left to do, these investments are relatively uninformed and therefore may often be redundant.

**CRL Action:** *CRL will expand the ICON database of newspaper holdings, and will enlist publishers of databases to contribute metadata on digitized materials to it as a service to the library community.*

- **Monitor the technology for production, management, and distribution of news content, to better understand how these systems work.** Information about the systems will be essential to libraries' ability to authenticate what is trustworthy for scholars, amid the ocean of content available that is of indeterminate origin. This forensic role will be important in evaluating potential "collecting" opportunities, as well as the cost of some digital content continues to rise.

**CRL Action:** *In 2011 CRL mapped the landscape of electronic news production for the Library of Congress Office of Strategic Initiatives, and will update that analysis periodically to keep pace with new technology and [developments](#).* ❖

**FOCUS on Global Resources**, published quarterly, is compiled by CRL's Communications Department. Gwen Ihnat, Virginia Kerr, Editors. Graphic design services provided by Molly O'Halloran, Inc. Special thanks to: Bernie Reilly, James Simon, Amy Wood, Kalev Leetaru, Debora Cheney, Connie Vinita Dowell, Joseph D. Combs, Sharon E. Farb, Todd Grappone, Megan Bernal, and Mark Phillips.

ISSN #: 0275-4924

### Center for Research Libraries Staff Contacts (800) 621-6044

President  
Bernard F. Reilly x 334  
breilly@crl.edu

Administrative Services Specialist  
Yvonne Jefferson x 319  
yjefferso@crl.edu

Member Liaison and Outreach  
Services Director  
Mary Wilke x 351  
mwilke@crl.edu

Director of International Resources  
James Simon x 324  
jsimon@crl.edu

Director of Technical Services  
Amy Wood x 327  
awood@crl.edu

Director of Information Systems  
Patricia Xia x 341  
pxia@crl.edu

Head, Access Services  
Kevin Wilks x 314  
kwilks@crl.edu

Head, Stack Management  
Patricia Finney x 328  
pfinney@crl.edu

Digital Program Manager  
Virginia Kerr x 265  
vkerr@crl.edu

Communications Specialist  
Gwen Ihnat x 289  
gihnat@crl.edu

### Global Resources Program Contacts (800) 621-6044

Director  
James Simon x 324  
jsimon@crl.edu

Project Coordinator—  
Global Resources Network  
Judy Alspach x 323  
jalspach@crl.edu

### Contact for Information Regarding: (800) 621-6044

Billing/Invoicing/Deposit Accounts  
Accounting Department x 345

Membership/Communications/  
Public Relations/Marketing  
Don Dyer x 317  
ddyer@crl.edu

CRL Publications Orders (non-*FOCUS*)  
Meeting/Schedules  
Yvonne Jefferson x 319

Loans/Photocopies/Demand  
Purchases  
Access Services Department x 314

Microfilm/Microfiche Sales  
Lorraine Farley x 336  
lfarley@crl.edu

Purchase Proposals/Electronic  
Licensing  
Mary Wilke x 351

Collection Deposits  
Mary Wilke x 351