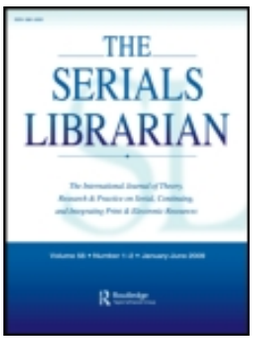

Routledge
Taylor & Francis Group

The Serials Librarian
From the Printed Page to the Digital Age

ISSN: 0361-526X (Print) 1541-1095 (Online) Journal homepage: https://www.tandfonline.com/loi/wser20

From a System of Journals to a Web of Objects

Herbert Van de Sompel & Susan Davis

To cite this article: Herbert Van de Sompel & Susan Davis (2015) From a System of Journals to a Web of Objects, The Serials Librarian, 68:1-4, 51-63, DOI: 10.1080/0361526X.2015.1026748

To link to this article: https://doi.org/10.1080/0361526X.2015.1026748

The Serials Librarian, 68:51–63, 2015
ISSN: 0361-526X print/1541-1095 online
DOI: 10.1080/0361526X.2015.1026748

Routledge
Taylor & Francis Group

From a System of Journals to a Web of Objects

HERBERT VAN DE SOMPEL
Presenter

SUSAN DAVIS
Recorder

Herbert Van de Sompel, Prototyping Team Leader, Research Library of the Los Alamos National Laboratory, took the audience on a journey from a paper-based scholarly communication system to a web of objects. Increasingly, the research process, not just its output, is web-based. Objects created and used as part of that process do not have the same sense of fixity that traditional publications had. Van de Sompel explored the problems inherent in the transition and the consequences of failing to proactively archive the full context of the scholarly record on the web.

KEYWORDS scholarly communication, web archiving, link rot, reference rot, web of objects

INTRODUCTION

Herbert Van de Sompel began his presentation by sharing an aspect of his background that is not mentioned on his personal website and established his credentials as a fellow serialist. Early in his career (about the mid/late 1980s), he was head of library automation at Ghent University in Belgium and collaborated with the University of Leuven in the Dortmund Library System-Leuven Library System (DOBIS-LIBIS) network to build an online catalog. Their first project was serials cataloging. For a year he worked a shift helping catalog serials for Ghent University Library. Unsurprisingly, this news was enthusiastically received by the NASIG audience. A lot has changed in the world of libraries and automation since then! Today, Van de Sompel and his research team are concerned with scholarly communication in the digital age. Everything they work on is about the transition of scholarly communication

from a paper-based system to a web-based system, which he likes to call the web of objects. By this he means a truly natively digital networked scholarly communication system that fully embraces the web and all its capabilities. This web of objects embraces aspects of the document web, the social web, and the Semantic Web. His team conducts research on this transition and its effects on the scholarly communication system.

CURRENT AND ONGOING PROJECTS

Van de Sompel described some of the projects he has worked on in the past. One of those projects was MEtrics from Scholarly Usage of Resources (MESUR), 2006–2008.[1] Once scholarly works appeared on the web, a wide array of metrics on impact could be gleaned that were not possible in a completely paper-based system. The focus of MESUR is on developing metrics from usage data by collecting massive amounts of information about downloads from publishers and library linking servers around the world. The project examined several questions. Is there a structure in the information? Can metrics be derived from that structure? They determined that the answer was yes. This work inspired and was a big catalyst for the altmetrics movement that has gained considerable attention over the past few years. This work also led to the bX recommender service (marketed by Ex Libris).

Memento (2000–2013) is described as time travel for the web.[2] Memento is an extension of the Hypertext Transfer Protocol (HTTP) and allows the user to request an older version of a web page, using that page's original uniform resource identifier (URI) and a preferred date in the past. Examples of use cases for this protocol are access to web archives and content management systems that support resource versioning such as wikis.

A current project that Van de Sompel's team in collaboration with the University of Edinburgh is working on is Hiberlink.[3] It is a two-year study funded by the Mellon Foundation to determine how many web links in online scholarly articles fail. Hiberlink will be discussed in more depth later in this article.

ResourceSync (2012–2014) is a project that was recently approved as ANSI/NISO Z39.99-2014 Standard.[4] Van de Sompel was a co-chair of the working group which developed this standard. The project revisits the problem domain of Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), which was released around 2001, but moves beyond metadata to include web resources in general.[5] The protocol is essentially about moving collections of information from location A to location B in a web-based manner.

The common denominator in all of his work is using the web to address scholarly communication issues. Van de Sompel is mainly concerned with information interoperability across systems, while achieving interoperability

by using web standards and web technologies. His focus is technical and targeted at machines, because he believes that increased accessibility for machines allows the creation of better tools for humans.

CORE FUNCTIONS OF THE SCHOLARLY COMMUNICATION SYSTEM

Van de Sompel placed his work in the context of fulfilling the core functions that characterize the scholarly communication system as identified by Roosendaal and Geurts:

- Register;
- Certify (Peer review);
- Awareness (Discoverability);
- Archive;
- Reward.[6]

All the projects he mentioned earlier relate to these functions in some way; for example, MESUR plays a role in the reward system; Memento and Hiberlink are concerned with archiving; and ResourceSync is about moving information from one place to another (e.g., registration systems to archival environments).

LONG-TERM ACCESS TO THE SCHOLARLY RECORD

For the remainder of his presentation Van de Sompel focused on one important aspect of the transition from a paper-based system to a truly native, web-based scholarly communication system: long-term access to the scholarly record. The prime question his team is investigating with this regard is: As scholarly communication evolves to a totally web-based system, will future scholars be able to revisit the scholarly record the way they could in the paper-based world? He offered several scenarios to illustrate the past, present, and future systems while issuing a warning about the future.

Paper-Based System

In the first scenario, Van de Sompel described the familiar paper-based system that has been around for hundreds of years, in which articles published in print journals reference articles in other print publications. Historically libraries have collected paper journals through subscriptions and archived them over the long term thereby providing researchers with the means to readily reconstruct the article and its references many years later. The

researcher is able to put together the entire context of the article and the cited articles, even if travel to several different libraries to obtain all the references was required.

Present System, Example One

In the second scenario, Van de Sompel provided two examples of the present. Articles in Portable Document Format (PDF) or Hyper Text Markup Language (HTML) format include links to referenced sources, not merely citations. All one has to do is click the link to retrieve the cited material. But who is archiving these online articles? For quite a while, it was uncertain who would tackle this role, and it was not necessarily libraries. Special purpose archival solutions emerged, such as Lots of Copies Keep Stuff Safe (LOCKSS), Portico, and e-Depot in the Netherlands. In this environment, how readily are researchers able to reconstruct the article and references a couple of years later? Not very likely without active intervention. A first reason is that links are brittle and may break. One reason why links may no longer work is because of publisher mergers and acquisitions which result in changes to the URIs, thus breaking the scholarly record. To solve this problem, persistent identifiers, such as the Digital Object Identifier (DOI), have emerged to redirect researchers to the current URI. The DOI functions as an intermediary and remains constant over time. Metadata about the resource is managed behind the scenes, where the DOI system maintains the current URI. CrossRef is an example of a system with a very successful infrastructure that provides some "interoperable glue" for the scholarly communication system. A second reason is that not enough electronic journal publications are actually archived. Optimistic estimates suggest that a maximum of 50% is archived in one of these special purpose archives, but the material that truly is in danger is not. David Rosenthal has written about preserving the scholarly record on his blog and in particular in a November 2013 blog post, "Patio Perspectives at ANADP II: Preserving the Other Half" citing a report from Luis Faria and other sources.[7,8,9] Rosenthal is greatly concerned about preservation of Open Access (OA) publications and other "high-risk" objects. Van de Sompel noted that the scholarly community does not have the means to accurately measure what is being archived; and therefore has no idea how much is at risk of being lost.

The Keepers Registry is a global monitor of archival arrangements for electronic journals that intends to provide an overview of electronic journals archived in special purpose archives.[10] While it is a promising development, it is still a work in progress and currently not able to provide an accurate overview. Van de Sompel is also deeply concerned that by basing its registry on International Standard Serial Number (ISSN), volume, and issue number, all of which are no longer relevant in the web world where URIs and DOIs reign, the Keepers Registry is living in the past.

Van de Sompel suggests that it should be possible to audit archival status of an article by means of its currency on the web, which is its URI and DOI. Such an audit is currently possible for regular resources on the web, but not for the journal literature. The Memento infrastructure, for example, allows polling web archives around the world to determine whether a page with a given URI is archived, and if so, within which timeframe. Using the URI for the Rosenthal blog for demonstration purposes, Van de Sompel showed that it is only archived in one place, the Internet Archive. In contrast, Memento found 429 archived copies of the NASIG website in the Internet Archive and in Archive.is. Back to the original question of reconstructing the context of the original publication in this incarnation, the answer is unclear, but probably no. One factor contributing to this negative conclusion is the uncertain and unverifiable archival status of electronic journal articles.

Present System, Example Two

The second example of the present system paints an even darker picture. Publications now contain links not just to other articles, but also to "web at-large" resources. These are resources that were needed or created during research activities such as project websites, workflows, blogs, ontologies, slides, data, software, and videos. These objects live on the web, are dynamic and ephemeral; they may change over time and are likely to disappear altogether. For example, consider how quickly software and hardware becomes obsolete. Data sets, videos, and blog may abruptly disappear. There is little assurance that the publication context can be reconstructed in such an environment; this is the problem domain at the core of the Hiberlink project. There is something that fundamentally distinguishes these kinds of resources from other resources such as the journal literature. They are not necessarily under the custodianship of parties that care about their long-term preservation (where the mindset is to finish the project and move on) in contrast to the large journal publishers who have a vested interest in long-term preservation of their content. These types of resources do not have the same sense of fixity because they, as previously discussed, frequently change, move or disappear. Hiberlink is researching this problem that they call "reference rot." Link rot is a term already used to describe links that stop working resulting in the dreaded 404 Not Found message. Reference rot is more than simply nonworking links. It contains another component called "content drift." Over time a referenced website may change and no longer be representative of what a researcher intended to link to. In the end, that link ends up worthless. Could the original context be reconstructed in this example? Definitely not.

Hiberlink's Research

Hiberlink's research strand works with large corpora of scientific information from 1997 through 2012. Electronic articles are automatically parsed and

URIs are extracted. URIs that link to the formal journal literature are filtered out based on the assumption that they are already being archived. Hence, Hiberlink's focus is on links to the web at-large, which are the ones at risk. Van de Sompel shared a chart that shows how articles are increasingly linking to the web at large; visually making the case for what is potentially vulnerable. For publication year 2012, there were over 120,000 links to web at-large resources in the PubMed Central corpus, whereas there were only a very few in 1997. Eighty percent of those 1997 links have stopped working; 15% to 20% of those from 2012 are rotten, and 30% to 40% from those around publication year 2006 do not work. This pattern is confirmed by other link rot studies using smaller samples.[11] Looking at the data from a different perspective, the percent of bad links compared to the number of references is not all that high, but it still represents a huge number of non-working links. For publication year 2012, the number is 20,000 dead links; projecting a link rot rate of 30% to 40% as determined by their research, this number will increase to 60,000 to 70,000 dead links by 2017. But, even if a link still works, say five years after the publication of the referencing paper, there is absolutely no guarantee that the content at the end of the link is still the same as it was at the time it appeared in the article. Quite to the contrary. Is the content of that URI the same as the time it was referenced? If not, then is there a way to locate a copy of that web page from the relevant time? Hiberlink's research to date clearly shows that there is a very serious need for web archives.

Hiberlink uses the Memento infrastructure to determine whether any of these web archives contain an archived version of a linked resource at the time it was referenced. Their research shows that the percent of resources that are archived is about the same as the level of link rot for recent content, 20%. By archived Van de Sompel means there exists a snapshot from within a month of the original time of reference captured in an archive. The numbers are quite dramatic. Eighty percent of the linked resources do not have an archived copy! Using the PubMed sample as the basis to extrapolate the total number of articles in the STM (science, technology, medicine) corpus, his team estimated that over a million articles from the 1997–2012 time period suffer from reference rot. Going back to the original research question of whether we can recreate the context of the original publication, the answer is a resounding no and something must be done to change that.

The previous analysis examined the STM literature, but the problem is not confined to STM. A recent article in the *New York Times*, "In Supreme Court Opinions, Web Links to Nowhere" by Adam Liptak, referenced a study at Harvard University that examined the law literature as well as Supreme Court decisions.[12,13] Their analysis showed that over 50% of the links in Supreme Court decisions exhibited reference rot. The reference rot problem exists outside of scholarly communication as well, where one familiar example is Wikipedia. At the end of the day, Van de Sompel explains that

the lack of web archives is a web at-large problem and the solution must also be in the web.

Proactive Approach to Web Archiving

Over the past ten years a small group of large institutions around the world have developed very significant tools for web archiving; however, more needs to be done than just web archiving in the way it is currently done. Typical web archiving follows a crawler-based paradigm, and, as a result the archived versions of referenced resources that are found in web archives are the result of *accidental* rather than on-demand archiving. In addition, the crawler-driven paradigm can result in reconstructed pages that are temporally incoherent. (For a more in-depth explanation see "A Framework for Evaluation of Composite Memento Temporal Coherence.")[14] Web pages consist of HTML and other embedded and linked content. Web crawlers might capture the HTML today, some of the embedded material tomorrow, and still other embedded content at a later date. Referring back to our problem statement of recreating the scholarly record, this crawling technique is inadequate because it does not recreate an accurate representation of the referenced website. Therefore, Van de Sompel promotes a pro-active, on-demand approach to web archiving. During the lifecycle of a scholarly paper, from note taking to publication, interventions should be made to actively archive referenced web resources. There are many opportunities to intervene. For example, the author could deposit the URI of websites visited while he or she is note taking and starting to write. Perma.cc is an example of a service that provides this functionality.[15] It was developed by the Harvard Law School Library in conjunction with a consortium of university law libraries around the world to tackle reference rot in the legal literature. The Hiberlink team at the University of Edinburgh has developed an experimental extension for the Zotero browser add-on that automatically pushes a web page into a web archive as the author bookmarks it. Van de Sompel and his team are working on the HiberActive tool that is aimed at intervening during manuscript or repository submission. HiberActive is notified when a new manuscript is submitted; it then pulls the manuscript across and parses out the referenced URIs; and next asks a web archive to archive the at-risk referenced URI. Once these referenced resources are archived, a user can use the Memento infrastructure to locate the temporally appropriate archived snapshots of a referenced resource. Using the NASIG homepage to demonstrate how Memento works, Van de Sompel searched for an archived version of the NASIG homepage for May 17, 2001. The Memento protocol was able to find an archived version of the homepage from May 19, 2001. He encouraged NASIG members to download the Memento plug-in for Chrome because it can be used to follow every link in time, or search his YouTube channel, hvdsomp@youtube, for two videos demonstrating how it works.[16]

In addition to the aforementioned approaches to web archiving during the lifecycle of an article, Van de Sompel encourages libraries to increase web archiving efforts with scholarly focus, by which he means resources that are likely to be referenced in scholarly communication. He suggests starting to archive institutional web pages and local research projects, and described several tools available today to use for archiving. There is an on demand subscription service called Archive-It, which is a subscription web archiving service from the Internet Archive to help organizations harvest, catalog, and manage their web collections.[17] His research team developed a tool called SiteStory that self-archives a webserver as it operates.[18] This tool has not been used as much as projected, perhaps because librarians did not understand how dire the situation has become.

Web Archives Do Not Last Forever

Van de Sompel convincingly demonstrated that the full context of the scholarly record cannot be recreated in the current environment, and there is an unintended consequence to depending on the URI from the web archive. The URI of the archived snapshot will work, but the original URI is its "currency" on the web and the way the resource is known on the web and in all the web archives around the world. Without the original URI, scholars will not be able to locate the resource in other archives, and thus become totally dependent on that one archive to access that copy. Alas, web archives are not guaranteed to last forever either; setting the stage to replace one link rot problem with another. Van de Sompel cited several examples of archives that have been in danger of fading away or disappeared altogether. WebCite, the first to offer web services to proactively archive websites as they were being referenced, experienced massive financial problems in 2013 and announced it needed to raise funds to continue operating (see WebCite article in Wikipedia).[19,20] Mummify is another example of a service that allowed users to make a permanent copy of anything online for free (after so many pages a paid subscription was required). Unfortunately the service only lasted a couple of months, and ironically a copy of its website in the Internet Archive is the only trace that Mummify ever existed.[21] Archive.is was the name of another archive until the service that manages the .is domains was attacked in mid-April 2014 and potentially jeopardized. Subsequently the manager of the archive renamed it archive.today. Had the .is top level domain indeed been compromised, then all links using the domain name archive.is would now be dead.

404 No More?

Realizing that web archives do not last forever and, hence that the links into web archives are as brittle as regular links, Van de Sompel and his team,

in collaboration with the Harvard Library Innovation Lab, Old Dominion University, and the Berkman Center for Internet & Society, were prompted to try to standardize a new way to link to archived resources that retains the original URI and augments the link with additional attributes to combat link rot and content drift.[22] This work has only recently begun and been dubbed the "404-No-More" project.[23] According to the project's website, the mset Attribute under development will provide authors with the means to provide a temporal reference to a URI, by augmenting the original link with additional attributes to document other copies as well as temporal information.

The Future Web of Objects

Over the course of his presentation, Van de Sompel has shown that recreating the publication context works well in the traditional paper-based system and does not work so well on the web in the two incarnations of the current journals system he described. However, the problems he identified in the current system are merely the tip of the iceberg. The future scholarly communication system—the web of objects as he refers to it—is where the real challenges emerge. While Van de Sompel does not purport to be a futurist, he listed three characterizations that are relevant to the discussion of recreating the scholarly context in the future:

1. The research process, not just its outcome, is becoming visible on the web;
2. There is an increased use of common web platforms for scholarship;
3. Communicated objects are very different than those in the paper-based system. They are heterogeneous, dynamic, compound, inter-related and all natively live on the web.

He listed a number of examples to illustrate these characteristics and show that the future is already here!

- Data, which has just exploded in the past five years to become an integral part of the web-based scholarly communication record;
- Software that may be core or auxiliary depending on the field, and if it is core it may be considered a scholarly contribution;
- Professional sites to share content, such as SlideShare for presentations and other content[24];
- Wikis are increasingly used for scientific scholarship, such as the WikiPathways for biology[25];
- NeuroLex, a neuroscience lexicon[26];
- Electronic lab notebooks (ELN) where you can document, share, and monitor an experiment;

- myExperiment (for bioinformatics) (http://www.myexperiment.org) is a social network for scientists to share scientific workflows and related materials.[27]

These examples ably illustrate those three characteristics of the future he outlined earlier and reiterated for emphasis. When the research process is visible it causes a massive extension of the scholarly record with new kinds of objects that are not easy to archive. There is an increased use of common platforms, such as GitHub, wikis, or SlideShare, that have many attractive features and give the appearance of archiving, but in reality they merely record the object.[28] Delving deeper in the terms and conditions of use for these platforms, rights are retained to remove content at any time. Trying to collect and archive those interrelated, heterogeneous, compound, and dynamic objects is not at all like archiving journal articles. We are presented with a new paradigm where it is important to collect all the interrelated pieces to preserve the entire object.

CONCLUSION

Wrapping up, Van de Sompel returned to the question he posed at the beginning of his presentation: whether it will be possible to recreate the original scholarly context in the future environment of the web of objects. He does not yet see an answer, and hopes his work with Hiberlink will start to build a perspective on the nature and magnitude of the problem. Two things are at play here. The first is the fundamental question of: What is the scholarly record? Where does it start, and where does it end? If the research process exists entirely on the web, what parts should be archived and which should not? Who will decide and how will the decision be made? The second question is: How can this be done? Proactive web archiving is only a small portion of the answer. And with those large, important questions resounding in our heads, Van de Sompel concluded his presentation.

QUESTIONS FROM THE AUDIENCE

What is the business model for a global system of sustainable archives as you've described? What are the likely costs on a global basis? What happens in the next economic downturn?

Van de Sompel has no real answer, and acknowledged that it is a really big problem to come up with a sustainable solution. It will be interesting to see if perma.cc can serve as a good model. It is presently a consortium of 50–60 law libraries, and he does not know what their business model is. Perhaps the LOCKSS/Controlled Lots Of Copies Keep Stuff Safe (CLOCKSS)

approach will be viable over a long period of time. Nonetheless, it remains a fundamental question that will drive how much and how long we can preserve scholarship.

Could you explain more about perma.cc?

Professor A visits a website that she feels is very important to reference in the paper she is writing and is concerned about the website's and/or the content's permanency. She does not want to end up with a useless URI and thus registers the URI of the website at perma.cc where it will be archived.

In the past, we heard that Brewster Kale with the Internet Archive was taking snapshots of the web and thus archiving the web.

This is exactly the reasoning leading to the "what happens if Brewster gets hit by a bus?" problem and leads back to his point that web archives do not last forever. Sustainability is not assured by a single web archive, even if it is the largest in the world. For many reasons, Van de Sompel believes there needs to be a distributed, coordinated effort for web archiving. For example, there was a political situation in the United Kingdom that resulted in the suppression from view of some speeches by David Cameron from the Internet Archive, permissible according to the Internet Archive's policy. Fortunately, copies existed elsewhere and the Memento architecture permitted access to these speeches, so the loss in one archive did not mean the speeches were completely lost.

With more and more mandates to deposit articles and other research products in institutional repositories (or other repositories), how do librarians motivate authors to prevent reference rot?

That's a very good question that relates to the social component of the problem. Hiberlink is looking at this a little bit. He and Peter Burnhill, Director of EDINA at Edinburgh University, hope to work with a group of researchers who should phenomenally care about the problem—PhD students. PhD candidates have a vested interest in this problem because the links in their theses and dissertations better work right when the review committees look at them! He agrees that just because tools exist, does not mean researchers will use them. Since researchers use bibliographic management software anyway, would it not be great if at the same time the software could seamlessly and automatically archive the references with little or no intervention by the researcher?

You previously talked about the issue that archiving a web page does not necessarily mean the embedded links are also archived. Can you share an example of a tool that both archives an object plus has a web crawler aspect?

This is why he insists on web archiving as a paradigm. Yes, many of the tools do this intrinsically. For example, the tool starts with a seed URI from

which you can ask it to follow links three generations deep. But where is the boundary, that is, in the web of objects how many levels deep do I need to delve to be sure I have all resources that make up the one object I am really interested in? It is similar to the splash page issue he described earlier. Machines are not able to determine what is relevant and what is not. He believes machine readable descriptions would help resolve this quandary. NASIG Past President, Bob Boissy, concluded the session by countering Van de Sompel's view that while he may not view himself as a futurist, he has provided NASIG with a great vision of the future.

NOTES

1. "MESUR: MEtrics from Scholarly Usage of Resources," Los Alamos National Laboratory, accessed June 11, 2014, http://mesur.lanl.gov/MESUR.html.
2. "Memento," accessed June 11, 2014, http://mementoWeb.org/.
3. "hiberlink," Los Alamos National Laboratory, accessed June 11, 2014, http://hiberlink.org/.
4. "ResourceSync (Resource Synchronization): A Joint NISO and Open Archives Initiative (OAI) Project," accessed May 23, 2014, http://www.niso.org/workrooms/resourcesync/.
5. "Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)," Protocol version 2.0 of 2002-06-14; document Version 2015-01-08, accessed May 23, 2014, http://www.openarchives.org/OAI/openarchivesprotocol.html.
6. Hans E. Roosendaal and Peter A. Th. M. Geurts, "Forces and Functions in Scientific Communication: An Analysis of Their Interplay," *Cooperative Research Information Systems in Physics*, received 7.12.1997; accepted 5.1.1998, accessed May 23, 2014, http://www.physik.uni-oldenburg.de/conferences/crisp97/roosendaal.html.
7. David Rosenthal, "DSHR's Blog," accessed May 20, 2014, http://blog.dshr.org.
8. David Rosenthal, "Patio Perspectives at ANADP II: Preserving the Other Half," *DSHR's Blog,* November 20, 2013, accessed May 20, 2014, http://blog.dshr.org/2013/11/patio-perspectives-at-anadp-ii.html.
9. Luis Faria et al., "Automatic Preservation Watch Using Information Extraction on the Web," accessed May 20, 2014, http://purl.pt/24107/1/iPres2013_PDF/Automatic%20Preservation%20Watch%20using%20Information%20Extraction%20on%20the%20Web.pdf.
10. "The Keepers Registry," accessed May 20, 2014, http://thekeepers.org
11. Sarah Rhodes, "Breaking Down Link Rot: The Chesapeake Project Legal Information Archive's Examination of URL Stability," *Law Library Journal* 102, no. 4 (Fall 2010): 581–597.
12. Adam Liptak, "In Supreme Court Opinions, Web Links to Nowhere," *New York Times*, September 23, 2013, accessed May 20, 2014, http://www.nytimes.com/2013/09/24/us/politics/in-supreme-court-opinions-clicks-that-lead-nowhere.html.
13. Jonathan Zittran, Kendra Albert, and Lawrence Lessig, "Perma: Scoping and Addressing the Problem of Link and Reference Rot in Legal Citations," *Harvard Public Law Working Paper* no. 13–42, accessed June 20, 2014, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2329161.
14. Scott G. Ainsworth, Michael L. Nelson, and Herbert Van de Sompel, "A Framework for Evaluation of Composite Memento Temporal Coherence," Cornell University Library, last revised October 5, 2014, http://arxiv.org/abs/1402.0928.
15. "Perma.cc," accessed June 23, 2014, http://perma.cc/about.
16. "Memento Time Travel," plug-in for Chrome, accessed June 27, 2014, http://bit.ly/memento-for-chrome.
17. "Archive.It," accessed May 20, 2014, https://www.archive-it.org.
18. "SiteStory," accessed June 27, 2014, http://mementoWeb.github.io/SiteStory/.
19. "WebCite," accessed June 23, 2014, http://www.Webcitation.org.
20. "WebCite," Wikipedia, last modified February 4, 2015, http://en.wikipedia.org/wiki/WebCite.
21. "Mummify," accessed June 23, 2014, https://Web.archive.org/Web/20140214233752/https://www.mummify.it/.

22. Ryan Westphal, Herbert Van de Sompel, and Michael L. Nelson, eds., "The mset Attribute," accessed June 23, 2014, http://berkmancenter.github.io/cache-link/mset-attribute.html.

23. Patrick Howell O'Neill, "This Project Aims to Make '404 Not Found' Pages a Thing of the Past," *The Daily Dot*, April 21, 2014, accessed June 24, 2014, http://www.dailydot.com/news/404-no-more-project/.

24. "SlideShare," accessed June 23, 2014, http://www.slideshare.net.

25. "WikiPathways," accessed June 23, 2014, http://wikipathways.org/index.php/WikiPathways.

26. "The NeuroLex Lexicon," Neuroscience Information Framework, accessed June 23, 2014, http://neurolex.org.

27. "myExperiment," accessed June 23, 2014, http://www.myexperiment.org.

28. "GitHub," accessed June 23, 2014, https://github.com.

CONTRIBUTOR NOTES

Herbert Van de Sompel is Prototyping Team Leader, Research Library of the Los Alamos National Laboratory, Los Alamos, New Mexico.

Susan Davis is Acquisitions Librarian for Continuing Resources, University at Buffalo, State University of New York, Buffalo, New York.