

# Journal of Contemporary Archival Studies

---

Volume 6

Article 23

---

2019

## Recognizing Co-Creators in Four Configurations: Critical Questions for Web Archiving

Amy Wickner  
[adw3@umd.edu](mailto:adw3@umd.edu)

Follow this and additional works at: <https://elischolar.library.yale.edu/jcas>



Part of the [Archival Science Commons](#)

---

### Recommended Citation

Wickner, Amy (2019) "Recognizing Co-Creators in Four Configurations: Critical Questions for Web Archiving," *Journal of Contemporary Archival Studies*: Vol. 6 , Article 23.  
Available at: <https://elischolar.library.yale.edu/jcas/vol6/iss1/23>

This Article is brought to you for free and open access by EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in *Journal of Contemporary Archival Studies* by an authorized editor of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact [elischolar@yale.edu](mailto:elischolar@yale.edu).

---

# Recognizing Co-Creators in Four Configurations: Critical Questions for Web Archiving

## **Cover Page Footnote**

As a content warning, this paper mentions violence in a university campus setting. Many thanks to the anonymous editors for their thoughtful comments, especially regarding structure. Thank you very much to Karl-Rainer Blumenthal, Samantha Abrams, Ed Summers, and Dusty Rhodes for reading early drafts of this work.

## Introduction

Web archives occupy a firm place in popular understandings of cultural heritage and memory. Web archiving is furthermore a core activity of many archives and other information institutions in the United States and internationally.<sup>1</sup> Amid ongoing research to develop best practices and new techniques for web archiving, one emerging area of study examines the sociotechnical factors that shape web archives and archiving. Contributing to this area of research, this paper explores how social forces, subjects, users, and technical agents co-create web archives and shape the everyday work of web archivists. Their impacts include all aspects of web archiving, from appraisal to description to access. I explore this research problem through four cases that support a critical approach to web archiving.

In the first section of the paper, I outline the research context by identifying four factors that co-create web archives with archivists. I then use four cases to illustrate how these co-creating factors interact in specific web archiving contexts. In the discussion, I synthesize observations from the cases and their implications for archivists. I conclude by noting areas for future work.

Key challenges for web archivists include recognizing the co-creators of web archives and deliberately engaging them in and through archival practice. Social contexts that influence web archives include social movements as well as government and corporate decisions that shape how people use the live web. Popular responses to evidence of institutional racism, tension between institutional memory and counter-memory, informal online memory practices, and the politics of archival education and labor also influence web archiving, in part by shaping appraisal practices. Users who shape web archives include researchers, participants in online creativity, and participants in archival education. Subject representation influences web archiving as archivists reproduce representations of subjects on the live web, choose content to save with respect for subjects, recognize gaps in web archives, acknowledge the ways subjects represent themselves, and account for their actions through honest description. Technical actors in web archiving include any and all web publishing, developing, and archiving technologies. In mapping the complexity of web archiving, this paper also proposes ways for archivists to confront and embrace this complexity in their work.

These observations and cases have several implications for archivists. For one, archivists must collaborate and build relationships with users and subjects of records that do not simply focus on collecting, archiving, and extracting content for preservation. Archivists must recognize repertoire and creative, interpretive reuse as forms of web archiving. Archivists must also ask what constitutes web archival repertoire, or transmitted practice. This paper contributes to research on the labor of web archiving by mapping the terrain in which web archivists learn and practice their craft. It suggests ways for web archivists to recognize the fullness of relationships in which they and their work exist, by looking beyond institutional and professional best practices.

It is possible to see the evidence presented in this paper as anecdotal and therefore limited in scope and relevance. As author, I am also an active subject in the cases. My specific biases and perspectives inevitably shape the argument. I name collaborators in order to give credit where it is due and emphasize that this paper offers one interpretation of shared experiences. Readers may find a lack of formal literature review or other conventions of research papers disorienting, although I try to provide a clear road map of the paper and highlight key points as I go. Acknowledging these limitations, I hope that this paper provides provocations and starting points for deeper empirical and theoretical investigations.

---

<sup>1</sup> Matthew Farrell, Edward McCain, Maria Praetzellis, Grace Thomas, and Paige Walker, *Web Archiving in the United States: A 2017 Survey* (Washington, DC: National Digital Stewardship Alliance, 2017), <https://osf.io/ht6ay/>; Miguel Costa, Daniel Gomes, and Mário J. Silva, "The Evolution of Web Archiving," *International Journal of Digital Libraries* 18, no. 3 (September 2017): 191–205, <https://doi.org/10.1007/s00799-016-0171-9>.

## Research context

The International Internet Preservation Consortium (IIPC) defines web archiving as “the process of collecting portions of the World Wide Web, preserving the collections in an archival format, and then serving the archives for access and use.”<sup>2</sup> This definition ably captures the objectives of web archiving as practiced in research libraries and archives. However, the complicated realities of web archiving can depart from this definition.

In this section, I establish four categories of co-creator that influence web archivists’ practices and shape web archives. Social contexts influence appraisal, or the analysis of social systems to understand the value of records. Users ask questions of and build upon web archives. Web archives represent subjects, often reproducing existing representations. Lastly, technical tools afford and limit web archiving activities.

### *Social forces shape web archives*

Web archiving is fundamentally an appraisal process. As the IIPC describes, “collecting portions” means not collecting the entirety of the web. Instead, archivists winnow and select within shifting sociotechnical conditions that include both systemic and individual biases. Archivists are not always able to articulate the fullness of criteria by which they understand archival value. Instead, colloquial rubrics of “saving” and institutional practices like collection development stand in for complex decision-making.

Appraisal refers to an array of processes through which archives and collections come to be. One major objective of appraisal is to analyze societal functions and translate the analyses into representative material documentation.<sup>3</sup> This means that broad social trends co-create web archives by shaping these archival analyses. Over the past decade, an increasing number of archivists and activists have begun to think deeply about how preserving web-based material can document aspects of social movements.<sup>4</sup>

Popular movements shape how people use the live web in highly visible ways. However, decisions by institutions, corporations, and other large organizations also have consequences for every other stakeholder in web archives. These bodies set the rules for what persists on the web. For example, federal policies like the 2017 repeal of net neutrality rules shape how people will be able to use the web in the near future.<sup>5</sup> Government information can disappear from and reappear on websites with little warning or explanation.<sup>6</sup> In November 2016, the Internet Archive announced that it would create a backup of its contents at several locations across Canada in response to the outcome of the US

---

<sup>2</sup> “Web Archiving,” International Internet Preservation Coalition, accessed November 19, 2018, <http://netpreserve.org/web-archiving/>.

<sup>3</sup> Terry Cook, “‘We Are What We Keep; We Keep What We Are’: Archival Appraisal Past, Present and Future,” *Journal of the Society of Archivists* 32, no. 2 (2011): 173–89, <http://dx.doi.org/10.1080/00379816.2011.619688>.

<sup>4</sup> See, for example: Bergis Jules, Ed Summers, and Vernon Mitchell Jr., *Ethical Considerations for Archiving Social Media Content Generated by Contemporary Social Movements: Challenges, Opportunities, and Recommendations*, April 2018, <https://www.docnow.io/docs/docnow-whitepaper-2018.pdf>; research presented at the National Forum on Ethics & Archiving the Web (March 22–24, 2018), <https://eaw.rhizome.org/>; and many others. Research and practice in this area draws on existing work by organizations like WITNESS (<https://witness.org/>), which focuses on video as a tool for human rights activism. I use “aspects of social movements” here to acknowledge the deep organizing that is less visible to outsiders but is crucial to growing and sustaining movements.

<sup>5</sup> Cecilia Kang, “F.C.C. Repeals Net Neutrality Rules,” *The New York Times*, December 14, 2017, <https://www.nytimes.com/2017/12/14/technology/net-neutrality-repeal-vote.html>.

<sup>6</sup> Richard Wiggins, “The Mysterious Disappearance of the White House Speech Archive,” *First Monday* 1, no. 2 (August 1996), <http://dx.doi.org/10.5210/fm.v1i2.472>; Chris Mooney and Juliet Eilperin, “EPA Website Removes Climate Science Site from Public View after Two Decades,” *The Washington Post*, April 29, 2017, <https://www.washingtonpost.com/news/energy-environment/wp/2017/04/28/epa-website-removes-climate-science-site-from-public-view-after-two-decades/>.

presidential election.<sup>7</sup> This example demonstrates that regime change affects web archives infrastructure.

The politics of business also determine whether information persists on the web. In November 2017, management at the online news outlets DNAinfo and Gothamist responded to workplace organizing by abruptly shutting down.<sup>8</sup> Journalists scrambled to recover copies of their clips. In November 2018, the photo sharing site Flickr notified users that it would limit free accounts to one thousand items and delete any data in excess of that limit in February 2019, beginning with the oldest images and videos.<sup>9</sup> On December 3, 2018, Tumblr banned and pledged to block NSFW (Not Safe for Work) content effective December 17.<sup>10</sup> In 2018, Reddit users noted that MySpace had lost all music uploaded to the site between 2003 and 2015 during a server migration.<sup>11</sup> The loss resurfaced in March 2019.<sup>12</sup> These examples highlight the vulnerability of digital content to top-down decision-making. Such dynamics must play a role in appraisal for web archives.

This section has argued that social forces like popular movements, government policy, and corporate decision-making shape web archives by affecting how people can use the web to share, access, and use information. Web archives practice must therefore account for these dynamics.

### *Users shape web archives*

Users and uses of live and archived websites co-create web archives with archivists. They do so by deploying archived sites as source material or evidence. Artists use them as “plural and heterogeneous archives” for remix.<sup>13</sup> Courts have begun to acknowledge web archives as legal evidence.<sup>14</sup> Social media users both cite and contribute material to web archives to use in online discussions.<sup>15</sup> Journalists report with web archives, gathering background and triangulating claims by public figures. Historians of the recent past use web archives and archiving to study political discourse and Internet histories, among other topics.<sup>16</sup> However, today’s web archives are so uncondusive to historical research

<sup>7</sup> Brewster Kahle, “Help Us Keep the Archive Free, Accessible, and Reader Private,” Internet Archive Blogs, November 29, 2016, <https://blog.archive.org/2016/11/29/help-us-keep-the-archive-free-accessible-and-private/>; Kahle, “FAQs about the Internet Archive Canada,” Internet Archive Blogs, December 3, 2016, <https://blog.archive.org/2016/12/03/faqs-about-the-internet-archive-canada/>.

<sup>8</sup> Andy Newman and John Leland, “DNAinfo and Gothamist Are Shut Down After Vote to Unionize,” *The New York Times*, November 2, 2017, <https://www.nytimes.com/2017/11/02/nyregion/dnainfo-gothamist-shutting-down.html>.

<sup>9</sup> Andrew Stadlen, “Why We’re Changing Flickr Free Accounts,” Flickr Blog, November 1, 2018, <https://blog.flickr.net/en/2018/11/01/changing-flickr-free-accounts-1000-photos/>.

<sup>10</sup> Shannon Liao, “Tumblr Will Ban All Adult Content on December 17th,” *The Verge*, December 8, 2018, <https://www.theverge.com/2018/12/3/18123752/tumblr-adult-content-porn-ban-date-explicit-changes-why-safe-mode>.

<sup>11</sup> JodiXD, “Myspace player won’t play songs, and I want to download them if possible,” Reddit, February 1, 2018, [https://www.reddit.com/r/techsupport/comments/7uiv8b/myspace\\_player\\_wont\\_play\\_songs\\_and\\_i\\_want\\_to/e3bfy1f/](https://www.reddit.com/r/techsupport/comments/7uiv8b/myspace_player_wont_play_songs_and_i_want_to/e3bfy1f/).

<sup>12</sup> BlueRajasmk2, “MySpace lost all music uploaded from 2003 to 2015,” Reddit, March 17, 2019, [https://www.reddit.com/r/technology/comments/b2381s/myspace\\_lost\\_all\\_music\\_uploaded\\_from\\_2003\\_to\\_2015/](https://www.reddit.com/r/technology/comments/b2381s/myspace_lost_all_music_uploaded_from_2003_to_2015/).

<sup>13</sup> Colin Post, “The Unending Lives of Net-Based Artworks: Web Archives, Browser Emulations, and New Conceptual Frameworks,” presented at the IIPC Web Archiving Conference, London, UK, June 16, 2017.

<sup>14</sup> Nicholas Taylor, “Understanding Legal Use Cases for Web Archives,” paper presented at the IIPC Web Archiving Conference, London, UK, June 16, 2017.

<sup>15</sup> Savvas Zannettou, Jeremy Blackburn, Emiliano De Cristofaro, Michael Sirivianos, and Gianluca Stringhini, “Understanding Web Archiving Services and Their (Mis)Use on Social Media,” ArXiv, January 2018, <https://arxiv.org/abs/1801.10396>.

<sup>16</sup> Megan Sapnar Ankerson, “Read/Write the Digital Archive: Strategies for Historical Web Research,” in *Digital Research Confidential: The Secrets of Studying Behavior Online*, ed. Eszter Hargittai and Christian Sandvig (Cambridge, MA: MIT Press, 2015), 29–54.

methods that it may not be accurate to refer to them as the “historical record.”<sup>17</sup> Archival processes of the past shape what questions can be asked and answered about the nineteenth and twentieth centuries. Similarly, web archiving practices will shape inquiry about the late twentieth century and beyond. Archivists working in institutions regularly rely on “research interest” as a factor in appraisal. The wide range of uses for web archives will shape web archiving into the future. Current and future users of web archives are key stakeholders in their development.

### ***Subject representation shapes web archives***

Archivists wield the power to name and classify. These forms of power have material effects on the world.<sup>18</sup> Access begins with archival description, which surfaces and contextualizes collections by linking them together. Archival description can conceal or make explicit the decisions and constraints in documentation that reveal archivists’ positions and name subjects.

Description and representation in web archives are specifically fraught because of how the live web represents subjects. Many people regularly encounter an Internet that is not built for them to use.<sup>19</sup> The web reproduces limited representations that impact subjects’ well-being and propagate harm.<sup>20</sup> Participatory spaces can actually marginalize difference. It can be difficult and draining to take back one’s identity, particularly amid discourses and policies that treat privacy protection as an individual rather than a collective responsibility.<sup>21</sup> Web archivists must recognize the power of archival representation to respond to the mis- and underrepresentation rampant on the web. Unless we deliberately imagine otherwise, web archivists will continue to build archives that perpetuate these unresolved issues. Furthermore, because anyone with a web presence is also represented in a web archive, web archivists are ourselves also subjects.

Representing subjects closely relates to access. This raises questions about consent and agency in preservation, as de Jesus argues. Describing a personal web archive of blogs by “marginalized people,” de Jesus says, “I believe they are important and ought to be preserved, but this doesn’t necessarily translate to unrestricted, public access.”<sup>22</sup> Archivists must consider when it’s best to unlink preservation and access in order to respect or protect subjects.

This section has argued that web archives can amplify harm in representing subjects. However, web archivists also wield the power to mitigate harm instead, for example by obtaining subjects’ informed consent to be documented.

---

<sup>17</sup> Susanne Belovari, “Historians and Web Archives,” *Archivaria* 83 (Spring 2017): 59–79, <https://archivaria.ca/index.php/archivaria/article/view/13600>; Ian Milligan, “Lost in the Infinite Archive: The Promise and Pitfalls of Web Archives,” *International Journal of Humanities and Arts Computing* 10, no. 1 (March 2016): 78–94, <https://doi.org/10.3366/ijhac.2016.0161>.

<sup>18</sup> Hope A. Olson, *The Power to Name: Locating the Limits of Subject Representation in Libraries* (Dordrecht: Springer, 2011); Geoffrey C. Bowker and Susan Leigh Star, *Sorting Things Out: Classification and Its Consequences* (Cambridge, MA: MIT Press, 1999).

<sup>19</sup> Maeve Duggan, *Online Harassment 2017* (Washington, DC: Pew Research Center, July 2017), <http://www.pewinternet.org/2017/07/11/online-harassment-2017/>.

<sup>20</sup> Alice Marwick and Rebecca Lewis, *Media Manipulation and Disinformation Online* (New York: Data & Society Research Institute, May 2017), <https://datasociety.net/output/media-manipulation-and-disinfo-online/>; Reporters without Borders, *Online Harassment of Journalists: Attack of the Trolls* (Paris: Reports without Borders, July 2018), [https://rsf.org/sites/default/files/rsf\\_report\\_on\\_online\\_harassment.pdf](https://rsf.org/sites/default/files/rsf_report_on_online_harassment.pdf).

<sup>21</sup> Amanda Lenhart, Michele Ybarra, Kathryn Zickuhr, and Myeshia Price-Feeney, *Online Harassment, Digital Abuse, and Cyberstalking in America* (New York: Data & Society Research Institute and Center for Innovative Public Health Research, November 2016), [https://www.datasociety.net/pubs/oh/Online\\_Harassment\\_2016.pdf](https://www.datasociety.net/pubs/oh/Online_Harassment_2016.pdf).

<sup>22</sup> Nina de Jesus, “Creating a Community-based Web Archive,” paper presented at the Gender and Sexuality in Information Studies Colloquium, Toronto, ON, October 18, 2014, <http://litwinbooks.com/2014papers/nina-de-jesus.pdf>.

### *Technical actors shape web archives*

The affordances of technical tools and platforms shape web content, web archiving, and access to web archives. Technical factors in the design of web technologies influence what one can archive. Trends in web design and development affect the archivability of websites. Designers increasingly use packaged pieces of code and dynamically loading content management systems.<sup>23</sup> Design decisions change browser technologies, as Google recently demonstrated by truncating Uniform Resource Locators (URLs) in its Chrome browser.<sup>24</sup> The personal technology market influences how people produce and consume web-based material. When the companies behind platforms change their terms of service, use of web trackers, or data selling policies, they influence how people use and archive the web. One example is changing what users can store, access, or collect via application programming interfaces (APIs).<sup>25</sup>

The sociotechnical systems of web archiving comprise not only people but also automated agents and technologies like bots, platforms, and infrastructure.<sup>26</sup> Distinct processes and outcomes result from using different technical methods for documenting the web, as the remainder of this section illustrates.

Web crawlers are one widely used web archiving method. A crawler begins to archive when a user specifies a starting “seed” URL. It creates and saves a facsimile of the seed, then identifies, follows, and copies links leading out from that page. The crawler repeats these steps until it reaches a user-specified limit defined in terms of host domain, number of documents, data, page quantity, or time. The Internet Archive maintains Heritrix, an open-source web crawler. Archive-It, a popular Internet Archive subscription service for managing institutional web archives and archiving, incorporates Heritrix among other technologies to perform captures. Archive-It captures appear in the Wayback Machine, the Internet Archive access portal for archived websites.

Emerging web archiving technologies do not make use of web crawlers in the same way. Ilya Kremer developed the browser-based software Webrecorder to capture dynamic, interactive elements of websites. Webrecorder records what takes place between browser and network as site content loads and a user interacts with it.<sup>27</sup> A Webrecorder user might click “record,” navigate around a site, and “replay” the resulting capture to check its fidelity. Users can draw from existing Internet archives and the live web to “patch” missing content. “Patching” means extracting pages from other archives or the live web. The underlying concept is to capture a specific experience involving a particular user, browser, and settings.

It is also possible to add URLs to an existing web archive, such as the Internet Archive or Archive.is. Perma.cc is a URL archiving service that emerged from concerns about link rot in legal opinions and

<sup>23</sup> Ashley Blewer, “Accessibility and Archivability,” BLOG Progress Process (blog), September 20, 2017, <http://bits.ashleyblewer.com/blog/2017/09/20/accessibility-and-archivability/>; Justin F. Brunelle, Mat Kelly, Michele C. Weigle, and Michael L. Nelson, “The Impact of JavaScript on Archivability,” *International Journal on Digital Libraries* 17, no. 2 (June 2016): 95–117, <https://dx.doi.org/10.1007/s00799-015-0140-8>.

<sup>24</sup> Lily Hay Newman, “Google Wants to Kill the URL,” *Wired*, September 4, 2018, <https://www.wired.com/story/google-wants-to-kill-the-url/>.

<sup>25</sup> Justin Littman, “Implications of Changes in Twitter’s Developer Policy,” Social Feed Manager Blog, May 18, 2017, <https://gwu-libraries.github.io/sfm-ui/posts/2017-05-18-twitter-policy-change>.

<sup>26</sup> Ed Summers and Ricardo Punzalan, “Bots, Seeds and People: Web Archives as Infrastructure,” in *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, ed. Charlotte P. Lee, Steve Poltrock, Louise Barkhuus, Marcos Borges, and Wendy Kellogg (New York: ACM, 2017): 821–34, <http://dx.doi.org/10.1145/2998181.2998345>.

<sup>27</sup> “About,” Webrecorder, accessed November 5, 2018, [https://webrecorder.io/\\_faq](https://webrecorder.io/_faq).

scholarly communication.<sup>28</sup> One can also use scripts to download files and data from web servers and online platforms. Youtube-dl, twarc, and wget are examples of such scripts.<sup>29</sup> Other emerging web archiving tools include peer-to-peer distributed file systems for web archives, browser extensions that produce Web ARChive (WARC) files, and standalone applications for replaying WARC files.<sup>30</sup> New web archiving technologies support different ways of thinking about preservation.

Just as users and uses shape web archives, web archives access systems shape use. For example, systems vary in how they “score” archived documents and present them for search and retrieval.<sup>31</sup> Access systems expose technical vulnerabilities, such as the ability to edit archived content.<sup>32</sup> This section has argued that technical agents shape web archives and archiving, from the creation of web content to capture, storage, and use.

### **Related work**

A critical position on web archiving includes archivists, social forces, users, subjects, and technical agents in its frame of analysis, considering how all of the above interact to co-create web archives. Emerging research seeks to understand the kind of work web archiving involves, who performs the work, and what perspectives they bring to the endeavor. Several recent efforts begin to map this research area. Through interviews with archivists and technologists who engage in web archiving, Summers and Punzalan identify technical and resource-related constraints that shape web archiving.<sup>33</sup> Post reviews research literature on appraisal in and of web archives, noting the potential to extend existing practices like collection policies and macro-appraisal.<sup>34</sup> Post also argues that cost-benefit analysis may be a useful way to assess the quantitative value of web archives and archiving. Maemura et al. further illuminate the work of web archiving through a study of three collections, their underlying workflows, and the role of “unexpected events” in web archiving at a university library.<sup>35</sup> In an ethnography of the Internet Archive, Ogden et al. make a case for “thick description” of how web

<sup>28</sup> Jonathan L. Zittrain, 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, SSRN, September 2013, <http://dx.doi.org/10.2139/ssrn.2329161>; Mia Massicotte and Kathleen Botter, “Reference Rot in the Repository: A Case Study of Electronic Theses and Dissertations (ETDs) in an Academic Library,” *Information Technology and Libraries* 36, no. 1 (March 2017): 11–28, <https://doi.org/10.6017/ital.v36i1.9598>.

<sup>29</sup> “GNU Wget,” GNU Operating System, updated September 15, 2017, <https://www.gnu.org/software/wget/>; “youtube-dl,” youtube-dl developers, accessed April 14, 2019, <https://rg3.github.io/youtube-dl/>; “twarc,” GitHub, DocNow, updated April 10, 2019, <https://github.com/DocNow/twarc>.

<sup>30</sup> “Software,” Web Science and Digital Libraries Research Group in the Department of Computer Science at Old Dominion University, modified June 19, 2018, <https://ws-dl.cs.odu.edu/Main/Software>.

<sup>31</sup> Thaeer Samar, Myriam C. Traub, Jacco van Ossenbruggen, Lynda Hardman, and Arjen P. de Vries, “Quantifying Retrieval Bias in Web Archive Search,” *International Journal of Digital Libraries* 19, no. 1 (March 2018): 57–75, <https://doi.org/10.1007/s00799-017-0215-9>.

<sup>32</sup> Ada Lerner, Tadayoshi Kohno, and Franziska Roesner, “Rewriting History: Changing the Archived Web from the Present,” in *CCS ’17: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, ed. Bhavani Thuraisingham, David Evans, Tal Malkin, and Dongyan Xu (New York: ACM, 2017), 1741–55, <https://doi.org/10.1145/3133956.3134042>.

<sup>33</sup> Summers and Punzalan, “Bots, Seeds, and People.”

<sup>34</sup> Colin Post, “Building a Living, Breathing Archive: A Review of Appraisal Theories and Approaches for Web Archives,” *Preservation, Digital Technology & Culture* 46, no. 2 (September 2017), <https://doi.org/10.1515/pdte-2016-0031>.

<sup>35</sup> Emily Maemura, Nicholas Worby, Ian Milligan, and Christoph Becker, “If These Crawls Could Talk: Studying and Documenting Web Archives Provenance,” *Journal of the Association for Information Science and Technology* 69, no. 10 (October 2018): 1223–33, <https://doi.org/10.1002/asi.24048>.



archivists conduct their work.<sup>36</sup> These sociotechnical approaches use theoretical foundations from archival studies, human-computer interaction, and practice theory to understand web archiving in context. In the current study, I lean heavily on evidence from personal experience as an archivist, student, and teacher, in response to these scholars' call for qualitative evidence of web archival practices.

## **Summary**

I have argued that the background shaping web archives includes social movements and institutions, current and future users of web archives, subjects represented on the web and in archives, and technical agents. Each background factor influences the others. For example, social forces shape the technical agents through which people use the web to represent subjects. As seen above, related work studies the "situated labor" of web archives and calls for further exploration of web archiving in practice. In the following section, I offer four cases that illustrate how social forces, users, subjects, and technical agents shape web archiving in four different contexts.

## **Shaping web archives: Four cases**

Building from this research context, this paper outlines an approach to web archiving through four cases that, weaving in and out of web archiving contexts, highlight opportunities to build upon and rethink current practices. My overall purpose is to model a kind of accounting through which web archivists might better understand their work and its impacts. The first case concerns collaborative appraisal to build a web archive documenting campus violence. The second case examines ways to read institutional discourses through gaps in web archives. The third case looks to online archives and creativity as informal memory practices. The fourth case describes an approach to teaching with web archives. I provide narrative context for each case and illustrate how it manifests the influence of social forces, users, subjects, and technical actors. I conclude each case by describing its implications for archivists.

In discussing these cases, I intend to give specific examples from personal experience that illustrate four factors influencing web archives and archiving. I describe dilemmas, decisions, and aspirations that I experience as an archivist doing web archiving. I also draw upon these cases to propose questions and ideas for future research and practice.

## **#FearTheTurtle**

The first case explores complexities in acknowledging and negotiating stakeholders in collaborative appraisal. The context is a web archive documenting a traumatic campus event.

On May 20, 2017, Lt. Richard W. Collins III, a senior at Bowie State University, a Historically Black University in Bowie, Maryland, was stabbed and killed by a stranger while visiting friends at the University of Maryland, College Park (UMD) campus.<sup>37</sup> The Prince George's County (Maryland)

---

<sup>36</sup> Jessica Ogden, Susan Halford, and Leslie Carr, "Observing Web Archives: The Case for an Ethnographic Study of Web Archiving," in *WebSci '17: Proceedings of the 2017 ACM on Web Science Conference*, ed. Peter Fox, Deborah McGuinness, Lindsay Poirer, Paolo Boldi, and Katharina Kinder-Kurlanda (New York: ACM, 2017), 299–308, <https://doi.org/10.1145/3091478.3091506>.

<sup>37</sup> Jessie Campisi and Evan Berkowitz, "UMD Student Charged in Connection with Stabbing near Montgomery Hall," *The Diamondback*, May 20, 2017, archived May 23, 2017, <https://wayback.archive-it.org/2410/20170523034926/http://www.dbknews.com/2017/05/20/university-police-investigate-homicide-regents-drive/>.

state's attorney's office will pursue hate crime charges for Collins's killer, a UMD student.<sup>38</sup> The state's attorney's office based this charge in part on digital evidence like smartphone data and social media activity. Extensive media coverage at the time of the murder reported that the accused had participated in white supremacist Facebook groups.<sup>39</sup>

### *Social forces: Students respond to systemic racism*

This tragedy only intensified existing conversations about how racism manifests on and around campus. In the days that followed, UMD students spoke on Twitter about Lt. Collins's death. They recalled numerous incidents throughout 2016 and 2017 in which members of the university community wrote, posted, and drew racial, ethnic, and ableist slurs and hate symbols in public places.<sup>40</sup> Students tweeted about harassment and micro-aggressions they had experienced from peers, instructors, administrators, and police. Some tagged their posts #FearTheTurtle to organize the dialogue around these topics. Social media users normally use this hashtag to rally around sports teams in honor of the UMD mascot, Testudo the Terrapin. The online discussion poignantly and pointedly co-opted athletics boosterism to voice personal stories about the violence of institutional racism.

### *Users of web archives*

A university web archive related to Richard Collins's murder and #FearTheTurtle came about thanks to engagement and interest from prospective users of archived websites.<sup>41</sup> At the UMD libraries, Special Collections and University Archives (SCUA) counts two interrelated digital humanities programs as neighbors: the Maryland Institute for Technology in the Humanities (MITH) and the program for African American History, Culture, and Digital Humanities (AADHUM). As conversations about campus climate developed over #FearTheTurtle, AADHUM project director Catherine Knight Steele expressed interest in collecting tweets from the hashtag. As a communication scholar, Dr. Steele wanted to document both social interactions and media coverage. These user interests provided an initial shape for the collection. As the SCUA archivist responsible for web archiving, I collaborated with Ed Summers, lead developer at MITH, to build a #FearTheTurtle web archive.

### *Representing subjects*

To respect the privacy of students tweeting about #FearTheTurtle and Richard Collins, the collection does not provide public access to tweets. Instead, it comprises breaking news stories and other articles that Twitter users shared widely in relation to UMD during a specific period of time. I supplemented these seeds with additional news coverage and commentaries—including articles from *The Diamondback*, UMD's largest student newspaper—and university statements. These additional URLs

---

<sup>38</sup> Lynh Bui, "U-Md. Student to Face Hate-Crime Charge in Fatal Stabbing on Campus," *The Washington Post*, October 17, 2017, archived October 18, 2017, [https://web.archive.org/web/20171018020842/https://www.washingtonpost.com/local/public-safety/u-md-student-to-face-hate-crime-charge-in-fatal-stabbing-on-campus/2017/10/17/a17bfa1c-b35c-11e7-be94-fabb0f1e9ffb\\_story.html](https://web.archive.org/web/20171018020842/https://www.washingtonpost.com/local/public-safety/u-md-student-to-face-hate-crime-charge-in-fatal-stabbing-on-campus/2017/10/17/a17bfa1c-b35c-11e7-be94-fabb0f1e9ffb_story.html); Bui, "Man Arrested in Fatal Stabbing at University of Maryland Will Not Face Federal Hate-Crime Charges," *The Baltimore Sun*, March 26, 2019, <https://www.baltimoresun.com/news/maryland/crime/bs-md-sean-urbanski-wont-face-federal-hate-crime-charges-0327-story.html>.

<sup>39</sup> Lynh Bui and Luz Lazo, "U-Md. Student Who Was Arrested in Fatal Stabbing to Appear in Court Monday," *The Washington Post*, May 22, 2017, archived May 22, 2017, [https://web.archive.org/web/20170522152653/https://www.washingtonpost.com/local/public-safety/u-md-student-arrested-in-fatal-stabbing-to-appear-in-court-today/2017/05/22/dc88cbb4-3eed-11e7-8c25-44d09ff5a4a8\\_story.html](https://web.archive.org/web/20170522152653/https://www.washingtonpost.com/local/public-safety/u-md-student-arrested-in-fatal-stabbing-to-appear-in-court-today/2017/05/22/dc88cbb4-3eed-11e7-8c25-44d09ff5a4a8_story.html).

<sup>40</sup> Andi Cwieka, "UMD Community Shares Experiences of Racism and Hate on Campus with #FearTheTurtle," *The Diamondback*, May 22, 2017, archived May 23, 2017, <http://wayback.archive-it.org/2410/20170523034912/http://www.dbknews.com/2017/05/22/umd-homicide-bowie-state-richard-collins-sean-urbanski-racism-feartheturtle/>.

<sup>41</sup> University of Maryland Libraries Web Resources Collection Program, "#feartheturtle," Archive-It, accessed July 20, 2018, <https://www.archive-it.org/collections/2410?fc=websiteGroup%3A%23feartheturtle>.

juxtapose media output with content created by the subjects that media coverage represents. Together, they outline the immediate context and campus climate surrounding the #FearTheTurtle discussion.

SCUA, MITH, and AADHUM have not solicited contributions to a more extensive web archive, but other organizations and individuals have created web-based material that responds to these events. Content creators include students, staff, faculty, and other community members of color. As we reach out to these creators, should we also build a web archive documenting hate speech specific to UMD? Should the archive reflect the perspectives of people who have remained oblivious to the roots and impacts of hate speech and other violence? Affording equal footing to perpetrators and victims of harm may in turn deepen the harm. It is also disingenuous to document expressions of hate as isolated “incidents,” “news,” or “events,” given that hate speech emerges from patterns of communication and behavior. Furthermore, its impacts are ongoing, everyday realities for the people most affected.

### *Technical actors*

In developing the collection, several human actors worked with and across a number of technical actors. Summers used a local instance of Social Feed Manager (SFM) to initiate a running collection of tweets that include the term “umd.” SFM is an “open source software that harvests social media data and web resources from Twitter, Tumblr, Flickr, and Sina Weibo.”<sup>42</sup> Data collection began several days after Lt. Collins’s death. However, in keeping with the terms of Twitter’s search API at the time, SFM also captured data for a sampling of tweets posted within seven days prior to the start of collection. Summers shared with me a list of URLs that most often appear in those tweets. Most were news articles. I added and crawled the URLs as “seeds” in SCUA’s Archive-It account. We intended to protect against link rot in the tweet dataset and to approximate a kind of community appraisal.

A combination of appraisal criteria and technical affordances shaped this part of the work. One aspect of web archives appraisal is “scoping” crawls. This means controlling a web crawler’s performance through data limits, document limits, regular expressions, and other technical rules that a user applies, for example through settings in the Archive-It interface. Another aspect is identifying significant properties of the web content. This informs the criteria for checking crawl quality. For example, I considered it important to capture embedded videos, which comprise the main news content of many URLs. Comment threads were also significant because they reflected the intensity of public conversation. However, the web crawler (Heritrix) struggled to capture and reproduce these dynamic elements. To preserve this material with more fidelity to the originals, I recorded the same group of URLs with Webrecorder. The ability to export WARC files from Webrecorder and upload them to Archive-It allowed me to unite, manage, and describe both groups of captures in one place.

### *Implications for archivists*

This case makes clear that documenting themes and events through online material is a relatively new activity for SCUA, although it is widespread and established in web archiving.<sup>43</sup> It highlights the value of thoughtfully and ethically navigating this documentary moment through relationships. For example, several memorial efforts to Lt. Collins are in progress as of this writing. Christian-Lamb has proposed a Richard Collins Community Archive that will engage a broad range of stakeholders in memory work

<sup>42</sup> Social Feed Manager website, George Washington University Libraries, accessed July 20, 2018, <https://gwu-libraries.github.io/sfm-ui/>.

<sup>43</sup> See, for example: “Spontaneous Event Collections,” Archive-It Blog, Archive-It, accessed November 4, 2018, <https://archive-it.org/blog/spontaneous-events/>; Roger Christman, “The Accidental Web Archive: The Tragedy at Virginia Tech Collection,” History @ Work (blog), National Council on Public History, March 14, 2018, <https://ncph.org/history-at-work/the-accidental-web-archive-the-tragedy-at-virginia-tech-collection/>.

related to racism on campus.<sup>44</sup> Early steps in this initiative include conversations with archival staff at SCUA and Bowie State University. The Critical Race Initiative, a research group based in the UMD Department of Sociology, recently launched a site called Justice 4 Richard Collins III to “remember Richard Collins III; aim to raise awareness around hate crimes; demand #Justice4RichardCollins; and affirm that #BlackLivesMatter.”<sup>45</sup> As well as supporting these initiatives, SCUA staff must continue a critical examination of our own practices. Acquiring or collecting material is not always an appropriate basis for building relationships and doing memory work.

### *Reading institutional discourses*

The second case examines how social forces, users and uses, subjects, and technical actors shape an institutional web archive’s capacity to reveal the life of the institution. Reflecting on #FearTheTurtle, I explored how UMD’s web archives might reflect aspects of the life of a campus.

#### *Social forces: The life of an institution*

SCUA’s current web archiving activities document official messaging. Online, institutional discourse exists in dialogue with #FearTheTurtle and other public conversations. It also interacts with university records that may not be visible even to archivists.

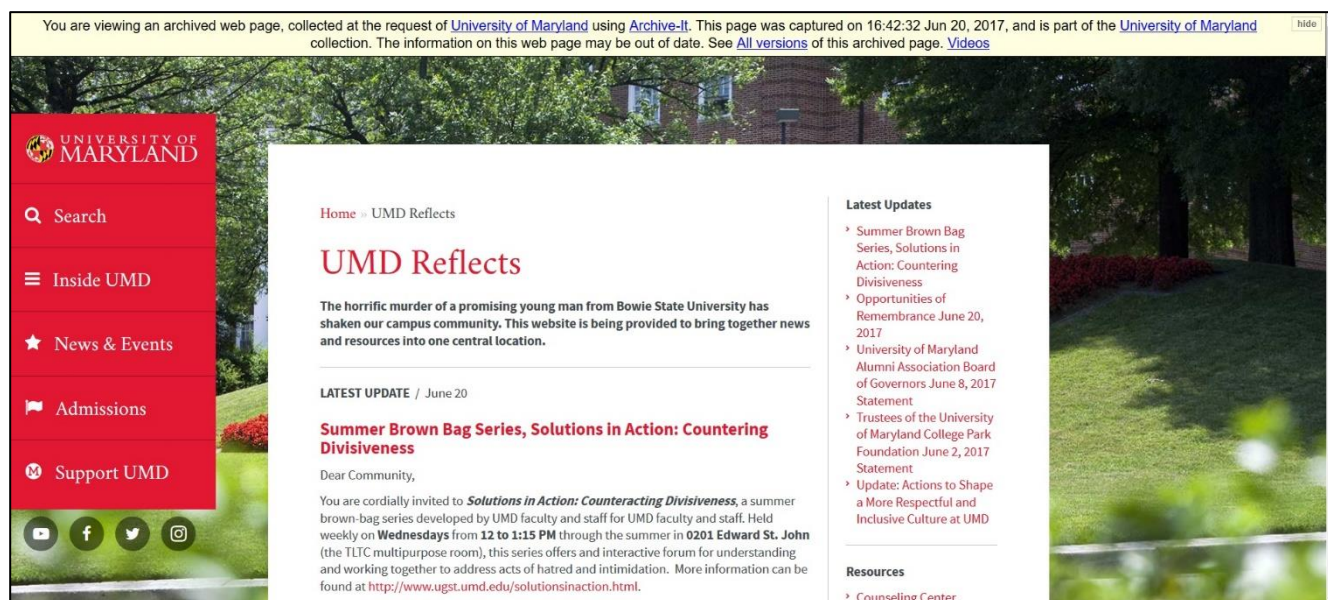


Figure 1. Archived snapshot of “UMD Reflects,” June 20, 2017.<sup>46</sup>

The past two years at UMD have seen public fora, meetings, communications, and miscommunication stemming from Richard Collins’s death and the campus community response. Since May 2017, university administrators have issued updates on their efforts to improve the campus climate using a website initially titled “UMD Reflects” (fig. 1). In August 2017, the president and University Senate

<sup>44</sup> Caitlin Christian-Lamb, “Remembrance and Representations of Hate: University Archives and Vernacular Memory of the Murder of Lt. Richard Collins III,” paper presented at ARCHIVES\*RECORDS 2018, Society of American Archivists/Council of State Archivists/National Association of Government Archivists and Records Administrators Joint Annual Meeting, Washington, DC, August 17, 2018.

<sup>45</sup> “About,” JUSTICE 4 RICHARD COLLINS III, accessed April 14, 2019, <https://justice4richardcollinsiii.weebly.com/about.html>.

<sup>46</sup> “UMD Reflects,” University of Maryland, archived June 20, 2017, <http://wayback.archive-it.org/2410/20170620164232/https://www.umd.edu/UMDreflects>.

appointed a task force on inclusion and respect, which formed working groups on free speech/hate speech, policies and procedures, prevention and education, climate, and hate/bias response.<sup>47</sup> One outcome of this work was a hate-bias protocol for incident reporting and investigation.<sup>48</sup> Between January and April 2018, the “UMD Reflects” site was renamed “We Are UMD” (fig. 2). This discursive move evokes unity in the aftermath of past conflict, eliding the continuity of campus tensions. This URL is part of the #FearTheTurtle seed group. In creating descriptive metadata, I have drawn attention to the rhetorical shift to explain use of a dual entry (“UMD Reflects/We Are UMD”) in the title field.

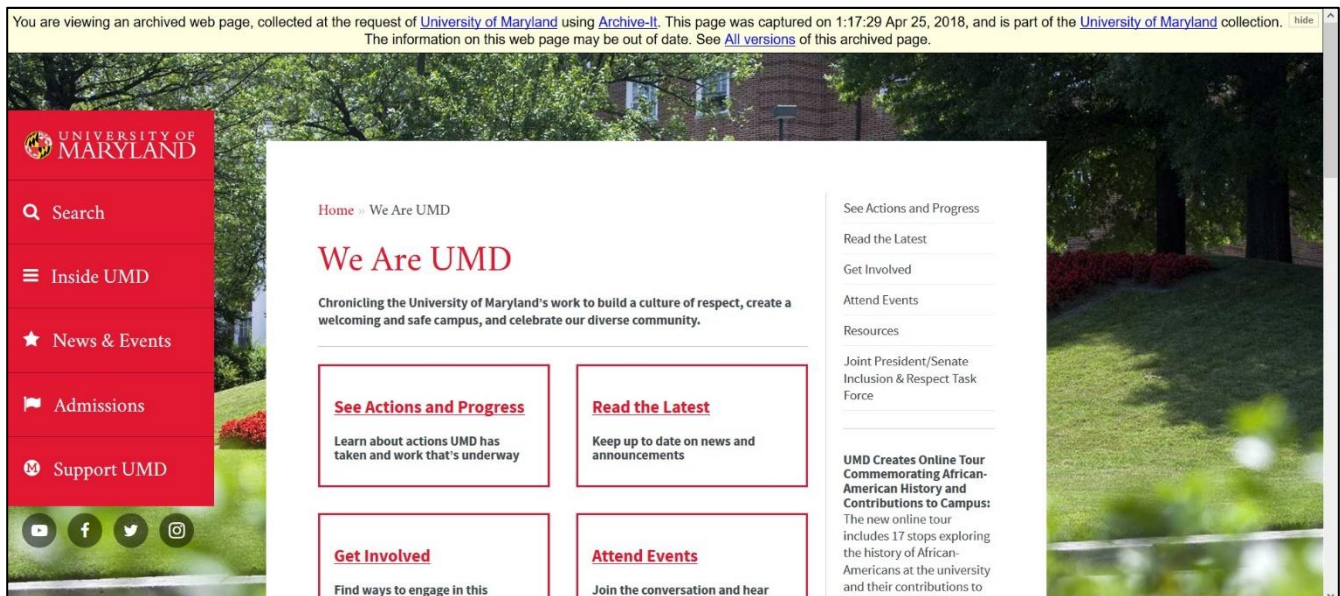


Figure 2. Archived snapshot of “We Are UMD,” April 25, 2018.<sup>49</sup>

### *Users of web archives*

Despite their institutional focus, these web archives are currently of little use for exploring university history. This is especially true without an understanding of how the campus, its departments, and websites are organized. Subject metadata exist at collection rather than seed or document level. Full-text, metadata, and URL search results can be voluminous and hard to parse. Users receive no guarantee that any given item exists to be found. Neither do web archives warn that material has not been preserved or is not accessible but perhaps should be. It may be more complicated than ever to understand the limits of what was collected, not to mention the difference between what was not documented and what never existed in the first place. In this specific institutional, political context, web archives obscure both institutional and archival decision-making.

<sup>47</sup> “Joint President/Senate Inclusion & Respect Task Force,” University of Maryland Senate, archived April 26, 2018, <http://wayback.archive-it.org/2410/20180426083138/https://senate.umd.edu/counciltaskforces/inclusion-respect>; “Inclusion and Respect at the University of Maryland,” University Senate Charge, University of Maryland Senate, August 25, 2017, archived May 1, 2018, [http://wayback.archive-it.org/2410/20180501124758/https://senate.umd.edu/system/files/resources/billDocuments/17-18-03/stage1/Inclusion\\_and\\_Respect\\_TF\\_Charge.pdf](http://wayback.archive-it.org/2410/20180501124758/https://senate.umd.edu/system/files/resources/billDocuments/17-18-03/stage1/Inclusion_and_Respect_TF_Charge.pdf); “Senate Bill 17-18-03,” University of Maryland Senate, archived April 29, 2018, <http://wayback.archive-it.org/2410/20180429040745/https://senate.umd.edu/searchBills/view?billId=621>.

<sup>48</sup> “Hate-Bias Incident Reporting Response Protocol,” University of Maryland Office of Diversity and Inclusion, archived April 25, 2018, <http://wayback.archive-it.org/2410/20180425090007/https://faculty.umd.edu/diversity/documents/hate-bias-protocol.pdf>.

<sup>49</sup> “We Are UMD,” University of Maryland, archived April 25, 2018, <http://wayback.archive-it.org/2410/20180424202411/https://umd.edu/weareumd>.



Web archives of official messaging support some forms of analysis, such as how these documents materialize through reference. Through them, university administrators equate policies or credentials with practices, citing commitment as evidence of tangible impact. Ahmed has identified and analyzed this pattern of behavior across institutions of higher education.<sup>50</sup> One might read a repository of official discourse to analyze how discourse does work in universities.

### *Representing subjects*

The ability to study institutional discourse does not sufficiently address concerns about representation in web archives, which I have described earlier in this paper. Official discourse creates specific roadblocks to representation. For example, critics of UMD argue that administrators are unable or unwilling to help students of color feel safe.<sup>51</sup> “Safe” here does not mean insulation from intellectual disagreement and argumentation—a common misunderstanding—but rather the visceral terms of not perpetually fearing for one’s physical and mental well-being. In light of these critiques, I went looking for ways that web archives represent the university’s past and present through unofficial perspectives.

UMD’s web archives include evidence of numerous protests, mostly through articles in *The Diamondback*, UMD’s student newspaper. People marched on the main administration building in 2005 to protest campus police use of pepper spray and night sticks, and in 2009 to support reinstating an assistant provost for equity and diversity.<sup>52</sup> A 2014 flyer (fig. 3) promotes a rally in response to police brutality. While interesting, these paltry search results barely represent decades of active and restless campus life.

### *Technical actors*

As well as underrepresenting certain subjects, web archives offer ample technical difficulty in studying the subjects they do represent. In the Wayback Machine, a header appears at the top of each archived page. It describes technical details of a given web crawl and directs visitors to all publicly available captures of a specific URL. These features can provide helpful context and navigation for anyone already familiar with web archives. However, user interface features do not support a new user of web archives whose area of inquiry is not tied to a specific URL.

Another area of technical difficulty lies in unexpected information management decisions by content creators. For example, as of a web crawl conducted on April 29, 2018, the document laying out the inclusion and respect task force charge had been moved to a new location in the Senate’s content management system.<sup>53</sup> The move broke one of the more visible links to this key document on the task force’s homepage. Broken links are obstacles for users of web archives and the live web alike, and mean difficult decisions for web archivists. For example, in a web archiving program staffed at less than 0.25 FTE, as mine is, systematically tracking down and recrawling specific pages is untenable.

<sup>50</sup> Sara Ahmed, *On Being Included: Racism and Diversity in Institutional Life* (Durham, NC: Duke University Press, 2012).

<sup>51</sup> Natalie Schwartz, “Students Call Out Admin with #UMDNotAHome after an Official’s Comments on Hate Symbol Bans,” *The Diamondback*, November 3, 2017, archived November 9, 2017, <https://web.archive.org/web/20171109025008/https://dbknews.com/2017/11/03/umd-twitter-hashtag-notahome-administration-hate-symbol-ban-free-speech/>.

<sup>52</sup> Kevin Litten, “Students Protest Against Racism,” *The Diamondback*, November 21, 2005, archived November 10, 2012, [https://wayback.archive-it.org/2410/20121110223144/http://www.diamondbackonline.com/article\\_1171ee0a-3d30-5b07-b11a-fab6dc646c3f.html](https://wayback.archive-it.org/2410/20121110223144/http://www.diamondbackonline.com/article_1171ee0a-3d30-5b07-b11a-fab6dc646c3f.html); Marissa Lang, “Students Demand Action, Explanation from Administrators,” *The Diamondback*, November 4, 2009, archived November 10, 2012, [https://wayback.archive-it.org/2410/20121110182302/http://www.diamondbackonline.com/news/campus/article\\_44e973ab-e40d-5e39-a71a-aea37ceba1ca.html](https://wayback.archive-it.org/2410/20121110182302/http://www.diamondbackonline.com/news/campus/article_44e973ab-e40d-5e39-a71a-aea37ceba1ca.html).

<sup>53</sup> “Page not found,” University of Maryland Senate, archived April 29, 2018, [http://wayback.archive-it.org/2410/20180429040731/https://senate.umd.edu/sites/default/files/resources/committeeFolders/InclusionRespectTF/Inclusion\\_and\\_Respect\\_TF\\_Charge.pdf](http://wayback.archive-it.org/2410/20180429040731/https://senate.umd.edu/sites/default/files/resources/committeeFolders/InclusionRespectTF/Inclusion_and_Respect_TF_Charge.pdf).



Figure 3. Archived website of WMUC, a student-run radio station at UMD, December 10, 2014.<sup>54</sup>

### *Implications for archivists*

One lesson from this case is that records management is a key concern for web archivists preserving institutional sites. Ineffective or unpredictable records management affects our ability to do this work. Another lesson is the need for web archives that document multiple angles and perspectives on administrative processes, including how governance responds to community positions and vice versa. One such community position is ProtectUMD, a list of demands for improved services and curriculum cooperatively developed and presented to the university by twenty-five student groups in November 2016.<sup>55</sup> Web archives can also show how institutions shift in response to national attention, such as an

<sup>54</sup> David Porter, "HIPHOP YOGA LIVE! POLICE PROTEST! NONVIOLENT THRU DOPE MUSIC!," WMUC College Park Radio 88.1FM, December 10, 2014, archived August 14, 2017, <https://wayback.archive-it.org/2410/20170814094121/http://wmuc.umd.edu/node/657>.

<sup>55</sup> Diamondback Staff, "ProtectUMD Has 64 Demands to Aid Vulnerable Students. These Are the Stories Behind Them," *The Diamondback*, December 13, 2016, archived December 13, 2016, <http://wayback.archive-it.org/2410/20161213034531/http://www.dbknews.com/protect-umd-demands/>.

ESPN investigative report on the May 29, 2018, death of UMD football player Jordan McNair.<sup>56</sup>

Building relationships that help to better document student life—as Hughes-Watkins proposes in a study of “reparative archives” in institutions<sup>57</sup>—can preserve perspectives in dialogue with official discourses. It is important to distinguish here between how students are heavily documented and data-mined<sup>58</sup> and how they represent themselves. Archivists often speak of “gaps in the record” or in collection holdings, but perhaps we would be better served by addressing gaps in our understanding. These anticipatory moves can create a framework for examining and pushing back against a variety of institutional discourses.

Indeed, to study how an institution works via its web history is simultaneously to study the web archives that act as primary sources. The limitations of UMD’s web archives require that users critically examine our appraisal and preservation practices. These practices closely connect to the institutional narratives the archives contain. Furthermore, the staffing model that attenuates web archives capacity stems from an institution-wide approach to supporting long-term programs with contingent labor.

### ***Web archiving repertoires***

As I discuss in the introduction, many institutions such as colleges, universities, government agencies, and cultural heritage organizations maintain web archives. However, non-institutional actors also have a stake in digital cultural heritage. The third case in this section explores how non-institutional memory work blurs boundaries among the creators, archivists, users, and subjects of online content. I use the following examples, neither comprehensive nor representative, as specific forms of online, collective memory work that complicate roles and other categories within web archives and archiving. This blurring makes it possible to imagine web archiving differently.

#### *Social forces: Online memory practice*

Informal memory practices take place through social media and other online communities. For example, many people pour labor into online creativity. It’s possible to interpret community-building practices in online archives as memory practices. In some communities, participants both use and maintain archives. De Kosnik demonstrates that members of online fan communities build archives through “repertoire,” or the repetition and transmission of archiving practice.<sup>59</sup> Repertoire encourages fellow fans and fan fiction readers to contribute content to the fan archives they frequent. Such content includes remixes of previous contributions. To reproduce a creative-archival practice in this way ensures its longevity, whether or not fan archives follow professional preservation best practices.

---

<sup>56</sup> Heather Dinich, Adam Rittenberg, and Tom VanHaaren, “The Inside Story of a Toxic Culture at Maryland Football,” *ESPN*, August 10, 2018, [http://www.espn.com/college-football/story/\\_/id/24342005/maryland-terrapins-football-culture-toxic-coach-dj-durkin](http://www.espn.com/college-football/story/_/id/24342005/maryland-terrapins-football-culture-toxic-coach-dj-durkin). For media coverage of campus responses, see, for example: Rick Maese, Roman Stubbs, and Nick Anderson, “‘A long, arduous task’: Inside the U-Md. Leadership Crisis,” *The Washington Post*, November 2, 2018, [https://www.washingtonpost.com/local/education/a-long-arduous-task-inside-the-u-md-leadership-crisis/2018/11/02/c9f5ff8c-deae-11e8-b732-3c72cbf131f2\\_story.html](https://www.washingtonpost.com/local/education/a-long-arduous-task-inside-the-u-md-leadership-crisis/2018/11/02/c9f5ff8c-deae-11e8-b732-3c72cbf131f2_story.html); Arya Hodjat, “Disillusioned with UMD’s Leadership, 24 Student Groups Rally to ‘Fire the Liars,’” *The Diamondback*, November 6, 2018, <http://www.dbknews.com/2018/11/06/maryland-football-protest-rally-fire-the-liars-wallace-loh-dj-durkin-jordan-mcnair-students-sga-2/>.

<sup>57</sup> Lae’l Hughes-Watkins, “Moving Toward a Reparative Archive: A Roadmap for a Holistic Approach to Disrupting Homogenous Histories in Academic Repositories and Creating Inclusive Spaces for Marginalized Voices,” *Journal of Contemporary Archival Studies* 5, article 6 (2018), <https://elischolar.library.yale.edu/jcas/vol5/iss1/6>.

<sup>58</sup> Kyle M. L. Jones and Dorothea Salo, “Learning Analytics and the Academic Library: Professional Ethics Commitments at a Crossroads,” *College & Research Libraries* 79, no. 3 (2018), <https://doi.org/10.5860/crl.79.3.304>.

<sup>59</sup> Abigail De Kosnik, *Rogue Archives: Digital Cultural Memory and Media Fandom* (Cambridge, MA: MIT Press, 2016).



### *Users and subjects of web archives*

Similarly, many online archives are actually databases or media-sharing accounts that rely on wide distribution as a form of preservation. One example is the Egyptian Mosireen YouTube channel, through which activists publish video evidence of police abuse.<sup>60</sup> Like the news and social media archives I discuss in the first part of this paper, this archive's availability hinges on a corporation's business model and content rules. While the Mosireen archive does not provide long-term, secure, geographically diverse storage or archival formats for its contents, it does ensure near-term access. Fast-tracking access to content on a widely familiar platform draws attention to the archive and the abuses it documents. Exposure may encourage others to contribute material. Like fan archives, this example illustrates the possibilities and limitations of archival repertoire and preservation through use. In publishing evidence of police abuse that has affected their own lives and communities, contributors to the channel play overlapping roles as archivists and the subjects represented in archives.

### *Representing subjects and users*

Archival repertoire and preservation through use blend the roles of creator, archivist, user, and subject. This blending foregrounds the creative, subjective nature of description and representation in archives. Two Twitter threads illustrate this. The sociologist and poet Eve Ewing initiated one thread on November 12, 2017.<sup>61</sup> Ewing began retelling notable Vines—short, looping videos that users created and circulated using the now defunct Vine app. Three examples follow:

“a girl dramatically waves her arm across a vast field of ducks. she speaks with pride.  
‘look at alllll these CHICKENS’”<sup>62</sup>

“A girl is asked several times by her mother to ‘do it for the Vine.’ She doesn’t want to.  
‘I ain’t gon do it,’ she says. At last, she relents.”<sup>63</sup>

“At a funeral service, a man dramatically tells the crowd that the deceased spirit is now freed, ‘LIKE DIS DOVE,’ which he releases into the air. The crowd reacts with audible horror when it is revealed that the dove is dead. Like super dead.”<sup>64</sup>

Other Twitter users recognized the Vines and responded with their own retellings. The thread continued for several hours on the hashtag #VinesWithoutVines. #VinesWithoutVines reanimated and celebrated precarious web content in a new form. Long-term preservation and access to original media were beside the point. Instead, participants in the hashtag enacted their simultaneous roles as creators,

<sup>60</sup> Amir-Hussein Radjy, “How to Save the Memories of the Egyptian Revolution,” *The Atlantic*, January 25, 2018, <https://www.theatlantic.com/international/archive/2018/01/an-internet-archive-rekindles-the-egyptian-revolutions-spirit/551489/>.

<sup>61</sup> Eve L. Ewing (@eveewing), “I should do a thread of my favorite vines but without links. Just me describing them from memory,” Twitter, November 12, 2017, <https://twitter.com/eveewing/status/929790289304346624>; Eve L. Ewing (@eveewing), “see guys we don’t need video to have fun with Vine on a rainy day!!!! we can use our imaginations,” Twitter, November 12, 2017, <https://twitter.com/eveewing/status/929790289304346624>.

<sup>62</sup> Eve L. Ewing (@eveewing), “a girl dramatically waves her arm across a vast field of ducks. she speaks with pride. ‘look at alllll these CHICKENS,’” Twitter, November 12, 2017, <https://twitter.com/eveewing/status/929793914835554306>.

<sup>63</sup> Eve L. Ewing (@eveewing), “A girl is asked several times by her mother to ‘do it for the Vine.’ She doesn’t want to. ‘I ain’t gon do it,’ she says. At last, she relents,” Twitter, November 12, 2017, <https://twitter.com/eveewing/status/929798988521406464>.

<sup>64</sup> Eve L. Ewing (@eveewing), “At a funeral service, a man dramatically tells the crowd that the deceased spirit is now freed, ‘LIKE DIS DOVE,’ which he releases into the air. The crowd reacts with audible horror when it is revealed that the dove is dead. Like super dead,” Twitter, November 12, 2017, <https://twitter.com/eveewing/status/929799380479172608>.

archivists, users, and subjects.

Filmmaker Matthew Cherry initiated another such thread on February 9, 2018, using the hashtag #GifHistory.<sup>65</sup> This thread featured reaction GIFs (Graphics Interchange Format): animated image files with which users of social media sites and group chat platforms make or illustrate points. Reaction GIFs often excerpt videos that circulate online. In the #GifHistory thread, Cherry matched extremely well-known reaction GIFs with their videos of origin. This form of web history reunites online content with its context. It invites reflection on the many uses and reinterpretations of reaction GIFs that take place in the intervening time. Creating and reusing reaction GIFs resembles the repertoire of fan archives that I discuss earlier in this section. Given its emphasis on context, #GifHistory also closely resembles traditional archival work like description or exhibit-making.

### *Technical actors*

Like fan archives and the Mosireen archive, #VinesWithoutVines and #GifHistory depend in part on technical actors. For example, Twitter's data retention models, timeline algorithm, and API specifications can support or discourage participation. How easy is it for users to stumble upon or resurface the relevant tweets? Twitter added the affordances that make these conversations possible over time: hashtags, threads, embedded GIFs, and video. #VinesWithoutVines came about because the Vine platform, Twitter-owned since 2012, sunsetted between 2016 and 2018.<sup>66</sup> These cases highlight relationships between subjects, users, and technical actors shaping web archives, and the social forces that influence their capacity to act.

### *Implications for archivists*

Online archives and threads like #VinesWithoutVines and #GifHistory expand the scope of what web archiving can be. They push back against equating archiving with long-term preservation. Web archiving can include interpretive memory practices that persist through archival repertoire and active, creative reuse. This section offers just a few examples of how specific attention to the characteristics of informal memory practices affords lessons for archival practice.

### *Critical curricula*

Imagining web archiving differently begins with supporting the growth of responsible, critical web archivists. It means taking care in introducing them to the technical and affective qualities of web archiving work. The following case reflects on an experience of teaching (with) web archiving, and how social forces, users, subjects, and technical actors interact in and out of the classroom.

### *Social forces: Archival education and appraisal*

In spring 2018, I taught with web archiving in a UMD College of Information Studies (iSchool) course called Documentation, Collection, and Appraisal of Records as co-instructor with Dr. Ricardo L. Punzalan. The course is an elective in the Master of Library and Information Science (MLIS) program's Archives and Digital Curation track. It addresses archival appraisal through four themed units: Appraisal Theories and Ideas (2 weeks), Institutional Appraisal Strategies (5 weeks), Non-Institutional Appraisal (2 weeks), and Ethics and Appraisal (5 weeks). We wove web archiving into the Ethics and Appraisal unit through five weekly themes: technology in appraisal; grief, loss, and affect; social justice; maintenance, labor, repair; and reappraisal and deaccessioning.

<sup>65</sup> Matthew A. Cherry (@MatthewACherry), "I wanna create a thread called #GifHistory. Send a gif that you want to know the backstory to and we'll try to find the original video," Twitter, February 9, 2018, <https://twitter.com/MatthewACherry/status/962011241815277568>.

<sup>66</sup> Ed Summers and Amy Wickner, "Archival Circulation on the Web: The Vine-Tweets Dataset," *Journal of Cultural Analytics*, Special Issue on Data Cultures, Culture as Data (June 4, 2019), <https://doi.org/10.31235/osf.io/6p8ea>.

We adapted the final, four-week assignment from a project that Ed Summers and Dr. Punzalan developed for the spring 2017 iteration of the course. We asked the class of twelve MLIS students<sup>67</sup> to develop a small documentation strategy for the March for Our Lives, a national, student-led demonstration against gun violence that took place on March 24, 2018. We asked students to approximate and appraise the universe of records creators around this topic. They identified key stakeholder communities and other organizations creating related documentation. Based on their findings, students identified websites, social media feeds, and other web-based material through which to build collections about specific aspects of the event.

Archive-It staff provided a data budget through an educational account associated with the iSchool. We asked students to run test crawls in Archive-It, examine crawl reports, adjust seed scopes, and describe their web archives using metadata fields in the administrative interface. They wrote appraisal justification essays and reflections on the process. We also discussed the assignment in class after it was due. This project foregrounded the analysis of social forces in the appraisal of web archives.

### *Users of web archives*

The students shaped web archives not only as appraising archivists but also as relatively new users of web archives and archiving technologies. Teaching with web archives and archiving makes a familiar medium strange. Websites and social media, like other everyday forms of digital reading and writing, can be easy to look “through” rather than “at”: “What is ubiquitous becomes transparent” until it breaks.<sup>68</sup> New web archivists can “break” everyday media as they poke and prod at how digital materials and environments are put together. The goal is to understand how web-based material could or should perform in a preserved state. Students work within unfamiliar technical constraints to interpret websites through the Archive-It interface and crawl reports. These experiential or affective sides of web archiving can be difficult to articulate. Instead, we opened these experiences for students to explore directly.

### *Representing subjects*

We encouraged students to write “honest description” for their web archives. Honest description reveals how archivists, custodians and intermediaries, and the process of archiving all shape archives.<sup>69</sup> It is a way to represent archival processes alongside the subjects of records. Users can then read web archives’ content in context. Many students laid out their appraisal decisions in metadata. Some addressed gaps in their collections and the reasons for those gaps. Many explained that they were new to web archiving and had operated under time constraints. By naming the conditions in which web archiving took place, students included archivists among the subjects of their web archives.

### *Technical actors*

As described above, the assignment asked students to “break” technologies in order to know them more intimately. Understandably, it proved difficult for many students to learn new tools and interfaces while also practicing, for nearly the first time, techniques like collaborative appraisal and integrating multiple understandings of archival value. One challenge for these new practitioner-theorists was evaluating the

---

<sup>67</sup> Many thanks to Emily Flint, Harrison Gage, Adam Gray, Eva Gunia, Eric Hung, Helen McNamara, Jen Piegols, Tracy Ritenour, Rosie Seidel, Jennifer Siegel, Andrew Staton, and Meaghan Wilson for their engagement with the course and this assignment.

<sup>68</sup> Christina Haas, *Writing Technology: Studies on the Materiality of Literacy* (Mahwah, NJ: L. Erlbaum Associates, 1996), xi.

<sup>69</sup> Jennifer Douglas, “Toward More Honest Description,” *The American Archivist* 79, no. 1 (Spring–Summer 2016): 26–55, <https://doi.org/10.17723/0360-9081.79.1.26>.

quality of their own work. That “all appraisal is local and subjective”<sup>70</sup> can sometimes feel like a lack of precedent for one’s actions.

### *Implications for archivists*

Just as fan archives persist through repertoire, hands-on and project-based archival education is a transmission of practice. As instructor, I attempted to model a specific web archiving repertoire and invite new archivists to join in. The project design inevitably reflected my goals, experiences, and assumptions as a student-teacher-archivist. As the above outcomes suggest, students extended this repertoire by walking a fine line between practiced judgment, technical constraints, and specific readings of archival theory. They also contributed critiques of the assignment by drawing on their personal experiences.

Students will interpret and transmit this praxis as they graduate and continue on in information work. For some, this repertoire will persist in archival spaces through appraisal, arrangement and description, policy development, documentation, mentorship, collegiality, advocacy, and teaching and learning. These are only a few of the highly interpretive activities that comprise archival work. However, today’s MLIS students are not guaranteed to find stable, well-paid jobs in archives, nor will they necessarily embrace the precarity that increasingly characterizes entry into the field.<sup>71</sup> As a result, repertoire transmitted in the graduate classroom, in relation to web archiving, may manifest in another context of information work altogether.

Teaching with web archiving has prodded me to build more training into my everyday work as digital archivist, which includes managing a web archiving program. I use training sessions and interviews to gather feedback from colleagues who experience technical constraints differently than I do. Their questions, observations, and distinct approaches to practice help to improve documentation such as training materials, manuals, and policies for a broad range of born-digital archival work.

## **Discussion**

I have presented four interrelated web archiving cases to demonstrate the configurations through which social forces and movements, users, subject representation, and technical actors shape web archives. I have also argued that these factors not only co-create web archives, but also influence the practice of web archiving. Throughout these cases, I note areas of friction between the intentions and realities of institutional web archiving. I also note areas for exploration and improvement. The following section synthesizes observations related to each influential factor and discusses implications for archivists in further detail. I revisit the four co-creators shaping web archives: social forces, users, subjects, and technical actors. I relate these co-creators to the four cases in the previous section to demonstrate that they also shape the practices of web archiving.

### ***Social forces shape web archiving***

As seen throughout these cases, a variety of social forces shape web archiving and archives. In an institutional context, university students responded not only to a traumatic act of violence but also to a

---

<sup>70</sup> Mark A. Greene and Todd J. Daniels-Howell, “Documentation with ‘An Attitude’: A Pragmatist’s Guide to the Selection and Acquisition of Modern Business Records,” in *The Records of American Business*, ed. James M. O’Toole (Chicago: Society of American Archivists, 1997), 162.

<sup>71</sup> Karly Wildenhaus, “Wages for Intern Work: Denormalizing Unpaid Positions in Archives and Libraries,” *Journal of Critical Library and Information Studies* 2, no. 1, Evidences, Implications, and Critical Interrogations of Neoliberalism in Information Studies (2019), <https://doi.org/10.24242/jclis.v2i1.88>; Angela Galvan, “Soliciting Performance, Hiding Bias: Whiteness and Librarianship,” *In the Library with the Lead Pipe* (June 3, 2015), <http://www.inthelibrarywiththeleadpipe.org/2015/soliciting-performance-hiding-bias-whiteness-and-librarianship>.

history of institutional racism. Their personal testimonies, published on social media, moved a professor, developer, and archivist to collaboratively appraise and build a web archive of media coverage and public statements. Elsewhere in the same institution, a history of campus unrest highlighted a tension in the web archives between archiving institutional records and documenting campus activity. I conclude from this case that these areas of documentation must exist in dialogue. Each form of documentation presents its own set of challenges to the resources and capacity of archivists. At best, the challenges suggest areas that archivists can target for outreach and personal growth.

Looking beyond institutional archives, the third case discusses informal memory practices, social media interaction, and online creativity. Although taking place outside of memory institutions, each of these practices has its own repertoire or set of conventions. One implication for archivists is the value of thinking beyond institutional best practices to recognize memory work. By understanding the purpose of such repertoires, archivists can better know how their archiving efforts contribute and are welcome, or do not and are not—linking back to de Jesus’s discussion of consent.

The graduate classroom is one site in which archivists transmit repertoire. In designing a curriculum that centers my own web archival repertoire, I transmit practices stamped with specific experiences and assumptions. These include my lived experience as a woman of color who entered the archival field through years of contingent employment, and my current access to resources as a full-time faculty archivist at a large institution. As students continue their careers as information workers, their experiences, resources, and interests will diverge. The politics of archival education and labor practices in archives are among the social forces shaping their web archiving repertoires.

### ***Users shape web archiving***

The four cases describe users’ influence on web archives. In one case, a communication scholar collaborated with archivists to appraise material she found to be of personal research interest. In another case, I identify gaps in web archives by imagining their potential use for researchers of institutional history. The case of online memory work describes archives, databases, and social media threads that are meant for insiders. Users or consumers of these memory sites also contribute to them by retelling, recontextualizing, or remixing. In the fourth case, multiple users interact with student-produced web archives. These user groups include hypothetical users of an institutional web archives, student peers, and co-instructors of the course. The expectations of each group of users shaped how students carried out a web archiving assignment.

### ***Subject representation shapes web archiving***

Each case highlights the influence of subjects and subject representation on web archives. In building a #FearTheTurtle web archive, colleagues and I chose to archive media coverage shared in tweets rather than the tweets themselves. We intended this ventriloquism to protect Twitter users’ privacy. However, we recognize the many subjects who remain unrepresented as a result. Future memory work on institutional racism must account for missing subjects and how best to represent them.

A case about institutional history suggests that web archives incompletely represent institutions as subjects. I identify a need to represent both institutions and their absent others. Documenting official discourse does not represent people working counter to or within institutions for change. Such subjects effect change in institutional practices and discourse over time, and will shape the course of web archiving one way or another.

Online memory practices reveal examples of how the subjects of web archives represent themselves and credit one another for creative contributions. As student and organizer Eva Dickerson noted at

Documenting Student Activism Now, social media represents just one facet of a person.<sup>72</sup> Similarly, viewing online memory practices as web archiving stresses that subjects are more complex than web crawling can capture. Archivists must acknowledge this by citing the incompleteness of web archival records. The case of students writing honest description suggests one approach to doing so. Archivists can also work closely with subjects to better represent their complexity—with fully informed consent, of course.

### ***Technical actors shape web archiving***

Throughout these cases, technical actors both enable and constrain web archiving. In building a web archives about #FearTheTurtle, archivists navigated the constraints of tware, Archive-It, and Webrecorder. In exploring web archives, users and archivists encounter broken links, which can result from information architecture changes. Memory practices outside of archives also rely on technical affordances. These include threading, content moderation, API configuration, and interface designs. Importantly, each technical actor represents an assemblage of human decision-making, itself under the influence of technical constraints. Similarly, new users of web archiving technology bring their own experiences to bear as they simultaneously learn concepts, interfaces, and ways to interpret data.

### ***Implications for archivists***

Collaboratively appraising a limited web archive of campus experiences suggests that web archives appraisal rests on relationships. Ethical web archiving involves forming and maintaining ethical relationships by thinking beyond the acquisition of content. Identifying gaps in web archives demonstrates that they stem from multiple causes. These include records management, archivists' (un)awareness of unofficial perspectives, and the overall lack of visibility of some voices relative to others. Known gaps point the way for web archivists to critique and rethink our work. Learning from online archives and social media threads, web archivists can better understand the memory practices and repertoires in use on the live web. Understanding web archiving as creative reuse can help web archivists acknowledge the interpretive nature of their work. In training memory workers of the future, web archivists transmit their archival repertoire(s) by openly negotiating relationships among users, subjects, and technical agents.

Although “web archiving” formally means preserving entire websites, we have seen through vernacular examples that the term may apply broadly to preserving any evidence of the web in use. An array of technical means support diverse interpretations of “web archivists” and “web archiving.” Web archiving increasingly requires not only competence with, but also the critical and curious use of tools in a growing landscape of technologies. Anyone teaching and learning with web archiving might focus their attention not only on how to use these technologies but also how they work.

Informal archiving practices support preservation through use and highlight the blended roles of archivists, subjects, and users of online archives. Even credentialed archivists take on multiple personae. Recognizing these overlapping forms of participation can help demystify web archiving for newcomers. It can also help archivists imagine web archiving differently by acknowledging complex relationships with those who have a stake in web archives. For example, developers and designers of web archiving systems also co-determine what is collected and how. Acknowledging and documenting this work contributes to the care of web archives. It makes more visible the infrastructures that support reading, writing, inquiry, and analysis with web archives.

As these cases demonstrate, web archiving involves archivists, subjects, users, technical agents,

---

<sup>72</sup> Eva Dickerson, panel discussion, Documenting Student Activism Now, Atlanta, GA, February 21, 2019, <https://standarchives.com/stand-symposium-auc-woodruff/>.

governments, corporations, communities, institutional structures, and alternative kinship ties as co-creators. In recognizing these co-creators or co-appraisers, archivists embody the potential to wrangle relational and generative power. However, they may not acknowledge or articulate this responsibility. To take critical ownership of web archiving, they must partner with those who know landscapes and subjects deeply. They must share with these partners what it is like to encounter web-based material with new eyes and through alternative technical constraints. Partners can also act as co-developers of curricula for navigating web archives. Such curricula address how web archives work, how they come to be, how to use them, where else they exist, and what is not in them.

## Conclusion

In this paper, I have argued that four categories of co-creator shape web archivists' practice and influence the development of web archives: social forces, users and uses, subjects, and technical agents. Four cases illustrate how these categories of co-creator overlap and interact in specific web archiving contexts. In these cases, social forces, users, subjects, and technical agents collaborate with archivists to shape documentation of the live web. I have recommended that web archivists acknowledge this complex array of contributors as a way to imagine web archives differently. A critical approach to web archiving recognizes relationships and blended roles among stakeholders, seeks opportunities for archival activity that do not center on acquisition, and acknowledges the value of creative reuse as an important aspect of preservation.

This paper contributes to ongoing research on the situated labor of web archiving through qualitative evidence of the questions, considerations, and dilemmas that web archivists encounter. Although writing from the perspective of an archivist embedded in an institution, I suggest that web archivists look beyond institutional best practices to better understand the social forces, users, subjects, and technical agents with which they co-create web archives. This recommendation carries several further implications for research and practice.

Archivists must encourage users of web archives to approach them as constructed. It is incumbent upon archivists to represent the mutual responsibilities that connect collectors, subjects, and users of web archives. Web archivists might begin by seeking collaborators who understand both the superficial and deep needs with which people approach web archives. What are users looking for? Do they care specifically about the web-ness of archives? How can web archives support and incorporate the knowledge, experience, and expectations that users bring to their inquiries?

Those who participate in web archiving can cultivate an open mind-set about what web archiving is and could be, who does it, and why. Recruiting the subjects of web archives into archiving practice could support more ethical collecting. It may also produce representation that acknowledges the impacts of web archiving and mitigates its potential to harm. Outreach efforts around web archiving must not always aim to "collect," but also recognize the importance of non-acquisitive, non-extractive interactions with members of a community. It is certainly important to ensure that people in the future can learn about a rich array of Internet lives. However, anyone springing into action to "save" online material must first consider whether a post or conversation is in any way theirs to preserve, whether it was intended to last, and what value they add by copying and recontextualizing it. The consenting participation of subjects represented in web archives is fundamental to this process. If choosing to proceed, web archivists must determine how best—and most honestly—to contextualize both the saved content and their own decisions. Through critical, reflexive study of the relationships involved in web archiving, web archivists can develop richer, more responsible approaches to practice.

## Bibliography

- “About.” Webrecorder. Accessed November 5, 2018. [https://webrecorder.io/\\_faq](https://webrecorder.io/_faq).
- Ahmed, Sara. *On Being Included: Racism and Diversity in Institutional Life*. Durham, NC: Duke University Press, 2012.
- Ankerson, Megan Sappan. “Read/Write the Digital Archive: Strategies for Historical Web Research.” In *Digital Research Confidential: The Secrets of Studying Behavior Online*, edited by Eszter Hargittai and Christian Sandvig, 29–54. Cambridge, MA: MIT Press, 2015.
- Archive-It. “Spontaneous Event Collections.” Archive-It Blog. Accessed November 4, 2018. <https://archive-it.org/blog/spontaneous-events/>.
- Belovari, Susanne. “Historians and Web Archives.” *Archivaria* 83 (Spring 2017): 59–79. <https://archivaria.ca/index.php/archivaria/article/view/13600>.
- Blewer, Ashley. “Accessibility and Archivability.” BLOG Progress Process (blog), September 20, 2017. <http://bits.ashleyblewer.com/blog/2017/09/20/accessibility-and-archivability/>.
- BlueRajasmayk2. “MySpace lost all music uploaded from 2003 to 2015.” Reddit, March 17, 2019. [https://www.reddit.com/r/technology/comments/b2381s/myspace\\_lost\\_all\\_music\\_uploaded\\_from\\_2003\\_to\\_2015/](https://www.reddit.com/r/technology/comments/b2381s/myspace_lost_all_music_uploaded_from_2003_to_2015/).
- Bowker, Geoffrey C., and Susan Leigh Star. *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT Press, 1999.
- Brunelle, Justin F., Mat Kelly, Michele C. Weigle, and Michael L. Nelson. “The Impact of JavaScript on Archivability.” *International Journal on Digital Libraries* 17, no. 2 (June 2016): 95–117. <https://dx.doi.org/10.1007/s00799-015-0140-8>.
- Bui, Lynh. “Man Arrested in Fatal Stabbing at University of Maryland Will Not Face Federal Hate-Crime Charges.” *The Baltimore Sun*, March 26, 2019. <https://www.baltimoresun.com/news/maryland/crime/bs-md-sean-urbanski-wont-face-federal-hate-crime-charges-0327-story.html>.
- . “U-Md. Student to Face Hate-Crime Charge in Fatal Stabbing on Campus.” *The Washington Post*, October 17, 2017. Archived October 18, 2017. [https://web.archive.org/web/20171018020842/https://www.washingtonpost.com/local/public-safety/u-md-student-to-face-hate-crime-charge-in-fatal-stabbing-on-campus/2017/10/17/a17bfa1c-b35c-11e7-be94-fabb0f1e9ffb\\_story.html](https://web.archive.org/web/20171018020842/https://www.washingtonpost.com/local/public-safety/u-md-student-to-face-hate-crime-charge-in-fatal-stabbing-on-campus/2017/10/17/a17bfa1c-b35c-11e7-be94-fabb0f1e9ffb_story.html).
- Bui, Lynh, and Luz Lazo. “U-Md. Student Who Was Arrested in Fatal Stabbing to Appear in Court Monday.” *The Washington Post*, May 22, 2017. Archived May 22, 2017. [https://web.archive.org/web/20170522152653/https://www.washingtonpost.com/local/public-safety/u-md-student-arrested-in-fatal-stabbing-to-appear-in-court-today/2017/05/22/dc88cbb4-3eed-11e7-8c25-44d09ff5a4a8\\_story.html](https://web.archive.org/web/20170522152653/https://www.washingtonpost.com/local/public-safety/u-md-student-arrested-in-fatal-stabbing-to-appear-in-court-today/2017/05/22/dc88cbb4-3eed-11e7-8c25-44d09ff5a4a8_story.html).
- Campisi, Jessie, and Evan Berkowitz. “UMD Student Charged in Connection with Stabbing near Montgomery Hall.” *The Diamondback*, May 20, 2017. Archived May 23, 2017. <https://wayback.archive-it.org/2410/20170523034926/http://www.dbknews.com/2017/05/20/university-police-investigate-homicide-regents-drive/>.
- Cherry, Matthew A. (@MatthewACherry). “I wanna create a thread called #GifHistory. Send a gif that you want to know the backstory to and we’ll try to find the original video.” Twitter, February 9, 2018. <https://twitter.com/MatthewACherry/status/962011241815277568>.
- Christian-Lamb, Caitlin. “Remembrance and Representations of Hate: University Archives and Vernacular Memory of the Murder of Lt. Richard Collins III.” Paper presented at ARCHIVES\*RECORDS 2018, Society of American Archivists/Council of State Archivists/National Association of Government Archivists and Records Administrators Joint Annual Meeting, Washington, DC, August 2018.



- Christman, Roger. "The Accidental Web Archive: The Tragedy at Virginia Tech Collection." History @ Work (blog), National Council on Public History, March 14, 2018. <https://ncph.org/history-at-work/the-accidental-web-archive-the-tragedy-at-virginia-tech-collection/>.
- Cook, Terry. "'We Are What We Keep; We Keep What We Are': Archival Appraisal Past, Present and Future." *Journal of the Society of Archivists* 32, no. 2 (2011): 173–89. <http://dx.doi.org/10.1080/00379816.2011.619688>.
- Costa, Miguel, Daniel Gomes, and Mário J. Silva. "The Evolution of Web Archiving." *International Journal of Digital Libraries* 18, no. 3 (September 2017): 191–205. <https://doi.org/10.1007/s00799-016-0171-9>.
- Cwieka, Andi. "UMD Community Shares Experiences of Racism and Hate on Campus with #FearTheTurtle." *The Diamondback*, May 22, 2017. Archived May 23, 2017. <http://wayback.archive-it.org/2410/20170523034912/http://www.dbknews.com/2017/05/22/umd-homicide-bowie-state-richard-collins-sean-urbanski-racism-feartheturtle/>.
- de Jesus, Nina. "Creating a Community-based Web Archive." Paper presented at the Gender and Sexuality in Information Studies Colloquium, Toronto, ON, October 2014. <http://litwinbooks.com/2014papers/nina-de-jesus.pdf>.
- De Kosnik, Abigail. *Rogue Archives: Digital Cultural Memory and Media Fandom*. Cambridge, MA: MIT Press, 2016.
- Diamondback Staff. "ProtectUMD Has 64 Demands to Aid Vulnerable Students. These Are the Stories Behind Them." *The Diamondback*, December 13, 2016. Archived December 13, 2016. <http://wayback.archive-it.org/2410/20161213034531/http://www.dbknews.com/protect-umd-demands/>.
- Dickerson, Eva. Panel discussion at Documenting Student Activism Now, Atlanta, GA, February 2019. <https://standarchives.com/stand-symposium-auc-woodruff/>.
- Dinich, Heather, Adam Rittenberg, and Tom VanHaaren. "The Inside Story of a Toxic Culture at Maryland Football." *ESPN*, August 10, 2018. [http://www.espn.com/college-football/story/\\_/id/24342005/maryland-terrapins-football-culture-toxic-coach-dj-durkin](http://www.espn.com/college-football/story/_/id/24342005/maryland-terrapins-football-culture-toxic-coach-dj-durkin).
- DocNow. "twarc." GitHub. Updated April 10, 2019. <https://github.com/DocNow/twarc>.
- Douglas, Jennifer. "Toward More Honest Description." *The American Archivist* 79, no. 1 (Spring–Summer 2016): 26–55. <https://doi.org/10.17723/0360-9081.79.1.26>.
- Duggan, Maeve. *Online Harassment 2017*. Washington, DC: Pew Research Center, July 2017. <http://www.pewinternet.org/2017/07/11/online-harassment-2017/>.
- Ewing, Eve L. (@eveewing). "a girl dramatically waves her arm across a vast field of ducks. she speaks with pride. 'look at alllll these CHICKENS.'" Twitter, November 12, 2017. <https://twitter.com/eveewing/status/929793914835554306>.
- . "A girl is asked several times by her mother to 'do it for the Vine.' She doesn't want to. 'I ain't gon do it,' she says. At last, she relents." Twitter, November 12, 2017. <https://twitter.com/eveewing/status/929798988521406464>.
- . "At a funeral service, a man dramatically tells the crowd that the deceased spirit is now freed, 'LIKE DIS DOVE,' which he releases into the air. The crowd reacts with audible horror when it is revealed that the dove is dead. Like super dead." Twitter, November 12, 2017. <https://twitter.com/eveewing/status/929799380479172608>.
- . "I should do a thread of my favorite vines but without links. Just me describing them from memory." Twitter, November 12, 2017. <https://twitter.com/eveewing/status/929790289304346624>.
- . "see guys we don't need video to have fun with Vine on a rainy day!!!!!! we can use our imaginations." Twitter, November 12, 2017. <https://twitter.com/eveewing/status/929793226025947136>.

- Farrell, Matthew, Edward McCain, Maria Praetzellis, Grace Thomas, and Paige Walker. *Web Archiving in the United States: A 2017 Survey*. Washington, DC: National Digital Stewardship Alliance, 2017. <https://osf.io/ht6ay/>.
- Galvan, Angela. "Soliciting Performance, Hiding Bias: Whiteness and Librarianship." *In the Library with the Lead Pipe* (June 3, 2015). <http://www.inthelibrarywiththeleadpipe.org/2015/soliciting-performance-hiding-bias-whiteness-and-librarianship>.
- "GNU Wget." GNU Operating System. Updated September 15, 2017. <https://www.gnu.org/software/wget/>.
- Greene, Mark A., and Todd J. Daniels-Howell. "Documentation with 'An Attitude': A Pragmatist's Guide to the Selection and Acquisition of Modern Business Records." In *The Records of American Business*, edited by James M. O'Toole, 161–229. Chicago: Society of American Archivists, 1997.
- Haas, Christina. *Writing Technology: Studies on the Materiality of Literacy*. Mahwah, NJ: L. Erlbaum Associates, 1996.
- Hess, Amanda. "Why Women Aren't Welcome on the Internet." *Pacific Standard*, January 6, 2014. <https://psmag.com/social-justice/women-arent-welcome-internet-72170>.
- Hodjat, Arya. "Disillusioned with UMD's Leadership, 24 Student Groups Rally to 'Fire the Liars.'" *The Diamondback*, November 6, 2018. <http://www.dbknews.com/2018/11/06/maryland-football-protest-rally-fire-the-liars-wallace-loh-dj-durkin-jordan-mcnair-students-sga-2/>.
- Hughes-Watkins, Lae'l. "Moving Toward a Reparative Archive: A Roadmap for a Holistic Approach to Disrupting Homogenous Histories in Academic Repositories and Creating Inclusive Spaces for Marginalized Voices." *Journal of Contemporary Archival Studies* 5, article 6 (2018). <https://elischolar.library.yale.edu/jcas/vol5/iss1/6>.
- JodiXD. "Myspace player won't play songs, and I want to download them if possible." Reddit, February 1, 2018. [https://www.reddit.com/r/techsupport/comments/7uiv8b/myspace\\_player\\_wont\\_play\\_songs\\_and\\_i\\_want\\_to/e3bfylf/](https://www.reddit.com/r/techsupport/comments/7uiv8b/myspace_player_wont_play_songs_and_i_want_to/e3bfylf/).
- "Joint President/Senate Inclusion & Respect Task Force." University of Maryland Senate. Archived April 26, 2018. <http://wayback.archive-it.org/2410/20180426083138/https://senate.umd.edu/councilstaskforces/inclusion-respect>.
- Jones, Kyle M. L., and Dorothea Salo. "Learning Analytics and the Academic Library: Professional Ethics Commitments at a Crossroads." *College & Research Libraries* 79, no. 3 (2018). <https://doi.org/10.5860/crl.79.3.304>.
- Jules, Bergis, Ed Summers, and Vernon Mitchell Jr. *Ethical Considerations for Archiving Social Media Content Generated by Contemporary Social Movements: Challenges, Opportunities, and Recommendations*. April 2018. <https://www.docnow.io/docs/docnow-whitepaper-2018.pdf>.
- JUSTICE 4 RICHARD COLLINS III. "About." Accessed April 14, 2019. <https://justice4richardcollinsiii.weebly.com/about.html>.
- Kahle, Brewster. "FAQs about the Internet Archive Canada." Internet Archive Blogs, December 3, 2016. <https://blog.archive.org/2016/12/03/faqs-about-the-internet-archive-canada/>.
- . "Help Us Keep the Archive Free, Accessible, and Reader Private." Internet Archive Blogs, November 29, 2016. <https://blog.archive.org/2016/11/29/help-us-keep-the-archive-free-accessible-and-private/>.
- Kang, Cecilia. "F.C.C. Repeals Net Neutrality Rules." *The New York Times*, December 14, 2017. <https://www.nytimes.com/2017/12/14/technology/net-neutrality-repel-vote.html>.
- Lang, Marissa. "Students Demand Action, Explanation from Administrators." *The Diamondback*, November 4, 2009. Archived November 10, 2012. [https://wayback.archive-it.org/2410/20121110182302/http://www.diamondbackonline.com/news/campus/article\\_44e973ab-e40d-5e39-a71a-aea37ceba1ca.html](https://wayback.archive-it.org/2410/20121110182302/http://www.diamondbackonline.com/news/campus/article_44e973ab-e40d-5e39-a71a-aea37ceba1ca.html).

- Lenhart, Amanda, Michele Ybarra, Kathryn Zickuhr, and Myeshia Price-Feeney. *Online Harassment, Digital Abuse, and Cyberstalking in America*. New York: Data & Society Research Institute and Center for Innovative Public Health Research, November 2016.  
[https://www.datasociety.net/pubs/oh/Online\\_Harassment\\_2016.pdf](https://www.datasociety.net/pubs/oh/Online_Harassment_2016.pdf).
- Lerner, Ada, Tadayoshi Kohno, and Franziska Roesner. "Rewriting History: Changing the Archived Web from the Present." In *CCS '17: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, edited by Bhavani Thuraisingham, David Evans, Tal Malkin, and Dongyan Xu, 1741–55. New York: ACM, 2017.  
<https://doi.org/10.1145/3133956.3134042>.
- Liao, Shannon. "Tumblr Will Ban All Adult Content on December 17th." *The Verge*, December 8, 2018. <https://www.theverge.com/2018/12/3/18123752/tumblr-adult-content-porn-ban-date-explicit-changes-why-safe-mode>.
- Litten, Kevin. "Students Protest against Racism." *The Diamondback*, November 21, 2005. Archived November 10, 2012. [https://wayback.archive-it.org/2410/20121110223144/http://www.diamondbackonline.com/article\\_1171ee0a-3d30-5b07-b11a-fab6dc646c3f.html](https://wayback.archive-it.org/2410/20121110223144/http://www.diamondbackonline.com/article_1171ee0a-3d30-5b07-b11a-fab6dc646c3f.html).
- Littman, Justin. "Implications of Changes in Twitter's Developer Policy." Social Feed Manager Blog, May 18, 2017. <https://gwu-libraries.github.io/sfm-ui/posts/2017-05-18-twitter-policy-change>.
- Maemura, Emily, Nicholas Worby, Ian Milligan, and Christoph Becker. "If These Crawls Could Talk: Studying and Documenting Web Archives Provenance." *Journal of the Association for Information Science and Technology* 69, no. 10 (October 2018): 1223–33.  
<https://doi.org/10.1002/asi.24048>.
- Maese, Rick, Roman Stubbs, and Nick Anderson. "'A long, arduous task': Inside the U-Md. Leadership Crisis." *The Washington Post*, November 2, 2018.  
[https://www.washingtonpost.com/local/education/a-long-arduous-task-inside-the-u-md-leadership-crisis/2018/11/02/c9f5ff8c-deae-11e8-b732-3c72cbf131f2\\_story.html](https://www.washingtonpost.com/local/education/a-long-arduous-task-inside-the-u-md-leadership-crisis/2018/11/02/c9f5ff8c-deae-11e8-b732-3c72cbf131f2_story.html).
- Marwick, Alice, and Rebecca Lewis. *Media Manipulation and Disinformation Online*. New York: Data & Society Research Institute, May 2017. <https://datasociety.net/output/media-manipulation-and-disinfo-online/>.
- Massicotte, Mia, and Kathleen Botter. "Reference Rot in the Repository: A Case Study of Electronic Theses and Dissertations (ETDs) in an Academic Library." *Information Technology and Libraries* 36, no. 1 (March 2017): 11–28. <https://doi.org/10.6017/ital.v36i1.9598>.
- Milligan, Ian. "Lost in the Infinite Archive: The Promise and Pitfalls of Web Archives." *International Journal of Humanities and Arts Computing* 10, no. 1 (March 2016): 78–94.  
<https://doi.org/10.3366/ijhac.2016.0161>.
- Mooney, Chris, and Juliet Eilperin. "EPA Website Removes Climate Science Site from Public View After Two Decades." *The Washington Post*, April 29, 2017.  
<https://www.washingtonpost.com/news/energy-environment/wp/2017/04/28/epa-website-removes-climate-science-site-from-public-view-after-two-decades/>.
- Newman, Andy, and John Leland. "DNAinfo and Gothamist Are Shut Down After Vote to Unionize." *The New York Times*, November 2, 2017.  
<https://www.nytimes.com/2017/11/02/nyregion/dnainfo-gothamist-shutting-down.html>.
- Newman, Lily Hay. "Google Wants to Kill the URL." *Wired*, September 4, 2018.  
<https://www.wired.com/story/google-wants-to-kill-the-url/>.
- Ogden, Jessica, Susan Halford, and Leslie Carr. "Observing Web Archives: The Case for an Ethnographic Study of Web Archiving." In *WebSci '17: Proceedings of the 2017 ACM on Web Science Conference*, edited by Peter Fox, Deborah McGuinness, Lindsay Poirer, Paolo Boldi, and Katharina Kinder-Kurlanda, 299–308. New York: ACM, 2017.  
<https://doi.org/10.1145/3091478.3091506>.

- Olson, Hope A. *The Power to Name: Locating the Limits of Subject Representation in Libraries*. Dordrecht: Springer, 2011.
- “Page Not Found.” University of Maryland Senate. Archived April 29, 2018. [http://wayback.archive-it.org/2410/20180429040731/https://senate.umd.edu/sites/default/files/resources/committeeFolders/InclusionRespectTF/Inclusion\\_and\\_Respect\\_TF\\_Charge.pdf](http://wayback.archive-it.org/2410/20180429040731/https://senate.umd.edu/sites/default/files/resources/committeeFolders/InclusionRespectTF/Inclusion_and_Respect_TF_Charge.pdf).
- Porter, David. “HIPHOP YOGA LIVE! POLICE PROTEST! NONVIOLENT THRU DOPE MUSIC!” WMUC College Park Radio 88.1FM, December 10, 2014. Archived August 14, 2017. <https://wayback.archive-it.org/2410/20170814094121/http://wmuc.umd.edu/node/657>.
- Post, Colin. “Building a Living, Breathing Archive: A Review of Appraisal Theories and Approaches for Web Archives.” *Preservation, Digital Technology & Culture* 46, no. 2 (September 2017). <https://doi.org/10.1515/pdtc-2016-0031>.
- . “The Unending Lives of Net-Based Artworks: Web Archives, Browser Emulations, and New Conceptual Frameworks.” Paper presented at the IIPC Web Archiving Conference, London, UK, June 2017.
- Radjy, Amir-Hussein. “How to Save the Memories of the Egyptian Revolution.” *The Atlantic*, January 25, 2018. <https://www.theatlantic.com/international/archive/2018/01/an-internet-archive-rekindles-the-egyptian-revolutions-spirit/551489/>.
- Reporters without Borders. *Online Harassment of Journalists: Attack of the Trolls*. Paris: Reporters without Borders, July 2018. [https://rsf.org/sites/default/files/rsf\\_report\\_on\\_online\\_harassment.pdf](https://rsf.org/sites/default/files/rsf_report_on_online_harassment.pdf).
- Samar, Thaer, Myriam C. Traub, Jacco van Ossenbruggen, Lynda Hardman, and Arjen P. de Vries. “Quantifying Retrieval Bias in Web Archive Search.” *International Journal of Digital Libraries* 19, no. 1 (March 2018): 57–75. <https://doi.org/10.1007/s00799-017-0215-9>.
- Schwartz, Natalie. “Students Call Out Admin with #UMDNotAHome After an Official’s Comments on Hate Symbol Bans.” *The Diamondback*, November 3, 2017. Archived November 9, 2017. <https://web.archive.org/web/20171109025008/https://dbknews.com/2017/11/03/umd-twitter-hashtag-notahome-administration-hate-symbol-ban-free-speech/>.
- Social Feed Manager website. George Washington University Libraries. Accessed July 20, 2018. <https://gwu-libraries.github.io/sfm-ui/>.
- “Software.” Web Science and Digital Libraries Research Group in the Department of Computer Science at Old Dominion University. Modified June 19, 2018. <https://ws-dl.cs.odu.edu/Main/Software>.
- Stadlen, Andrew. “Why We’re Changing Flickr Free Accounts.” Flickr Blog, November 1, 2018. <https://blog.flickr.net/en/2018/11/01/changing-flickr-free-accounts-1000-photos/>.
- Summers, Ed, and Amy Wickner. “Archival Circulation on the Web: The Vine-Tweets Dataset.” *Journal of Cultural Analytics*, Special Issue on Data Cultures, Culture as Data (June 4, 2019). <https://doi.org/10.31235/osf.io/6p8ea>.
- Summers, Ed, and Ricardo Punzalan. “Bots, Seeds and People: Web Archives as Infrastructure.” In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, edited by Charlotte P. Lee, Steve Poltrock, Louise Barkhuus, Marcos Borges, and Wendy Kellogg, 821–34. New York: ACM, 2017. <http://dx.doi.org/10.1145/2998181.2998345>.
- Taylor, Nicholas. “Understanding Legal Use Cases for Web Archives.” Paper presented at the IIPC Web Archiving Conference, London, UK, June 2017.
- “UMD Reflects.” University of Maryland. Archived June 20, 2017. <http://wayback.archive-it.org/2410/20170620164232/https://www.umd.edu/UMDreflects>.
- University of Maryland Libraries Web Resources Collection Program. “#feartheturtle.” Archive-It. Accessed July 20, 2018. <https://www.archive-it.org/collections/2410?fc=websiteGroup%3A%23feartheturtle>.

- University of Maryland Office of Diversity and Inclusion. *Hate-Bias Incident Reporting Response Protocol*. Archived April 25, 2018. <http://wayback.archive-it.org/2410/20180425090007/https://faculty.umd.edu/diversity/documents/hate-bias-protocol.pdf>.
- University of Maryland Senate. *Inclusion and Respect at the University of Maryland*. August 25, 2017. Archived May 1, 2018. [http://wayback.archive-it.org/2410/20180501124758/https://senate.umd.edu/system/files/resources/billDocuments/17-18-03/stage1/Inclusion\\_and\\_Respect\\_TF\\_Charge.pdf](http://wayback.archive-it.org/2410/20180501124758/https://senate.umd.edu/system/files/resources/billDocuments/17-18-03/stage1/Inclusion_and_Respect_TF_Charge.pdf).
- . *Senate Bill 17-18-03*. Archived April 29, 2018. <http://wayback.archive-it.org/2410/20180429040745/https://senate.umd.edu/searchBills/view?billId=621>.
- “We Are UMD.” University of Maryland. Archived April 24, 2018. <http://wayback.archive-it.org/2410/20180424202411/https://umd.edu/weareumd>.
- “Web Archiving.” International Internet Preservation Coalition. Accessed November 19, 2018. <http://netpreserve.org/web-archiving/>.
- Wiggins, Richard. “The Mysterious Disappearance of the White House Speech Archive.” *First Monday* 1, no. 2 (August 1996). <http://dx.doi.org/10.5210/fm.v1i2.472>.
- Wildenhaus, Karly. “Wages for Intern Work: Denormalizing Unpaid Positions in Archives and Libraries.” *Journal of Critical Library and Information Studies* 2, no. 1, Evidences, Implications, and Critical Interrogations of Neoliberalism in Information Studies (2019). <https://doi.org/10.24242/jclis.v2i1.88>.
- “youtube-dl.” youtube-dl developers. Accessed April 14, 2019. <https://rg3.github.io/youtube-dl/>.
- Zannettou, Savvas, Jeremy Blackburn, Emiliano De Cristofaro, Michael Sirivianos, and Gianluca Stringhini. “Understanding Web Archiving Services and Their (Mis)Use on Social Media.” Preprint, last revised April 9, 2018. <https://arxiv.org/abs/1801.10396>.
- Zittrain, Jonathan L., 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, SSRN, September 2013. <http://dx.doi.org/10.2139/ssrn.2329161>.