

Lorena Ramírez-López
December 12, 2014
Digital Preservation
Professor: Kara Van Malssen

Usability and Web Archiving

User Experience in Archive-It



It is not difficult to take the Internet for granted. With its immediate gratification and global access, the World Wide Web is a unique resource for information that our society heavily relies on daily. It is a resource that not only generates thousands of giga-to petabytes of information, but also loses thousands of digital libraries each day as data on the web simply disappears (IIPC). With the high amount of legal responsibilities and limited amount of server space and time, archiving the web seems like an impossible task. For the most part, it can be extremely difficult, which is why services, such as Archive-It that collect and access cultural heritage on the web, are important because they facilitate in the web archiving process. However, organizations only began archiving the Internet in 1996 and creating tools about a decade ago when the Internet Archive started the Heritrix web crawler development back in 2003. The idea of actual web archiving services is even newer. It is an issue that should be addressed when talking about Archive-It and other services and about how useful and productive the service as a tool can be and what can be done to make the services more usable for web archiving users.

“Web archiving is 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 (IIPC).” Similarly to other materials and information that libraries and archives collect, web archiving content from the Internet must be selected, harvested, and preserved so a user can ultimately access it. “Regardless of its form, there are many reasons why we archive information...[fundamentally it] is the recognition that an information object holds value outside the original purpose for which it was created...(Penncock 3).” Today, members of the International Internet Preservation Consortium, which is a membership organization dedicated to improving the tools, standards and best practices of web archiving, strongly argue that the web is our generation’s human record and it must be preserved now for the use and access of future generations (IIPC).

One of the main reasons for archiving websites, particularly in the cultural heritage community, is the relatively short timeframe within which content can be, and has been, ‘lost’: the average lifespan of web pages...range from 44 days to 75 or 100. Specific content on a page can disappear even more frequently, particularly on news-driven and social media websites (Penncock 3). There are also legal obligations required by some institutions to capture records for governmental officials and researchers. And then there is general social interest on what’s happening on the web. “The web archiving question should be viewed in the perspective of [cultural] heritage preservation...by nature, Web content is ephemeral: sites disappear continually and are frequently updated, involving the disappearance of what often is very valuable online information and Web content. [Therefore] archiving the Web requires a special attention in order to retain its value and ensure its greater fidelity (IMF).” It is these inherent reasons on why we should archive the web that are also the inherent challenges when archiving the web. The challenge of capturing a vast amount of data in a short amount of time is not only determined by the content’s ephemerality, but also “the alleged immensity of the Web, where storage costs and capacity of automatic tools to gather huge amount of information need to be considered...(Masanes 8).” Selection policies for web archives are determined by broader, organizational collection policies that are scoped based on domain collections

and selective collections, however, by establishing scoped collections it creates an artificial limit that the Internet (in general) does not abide by- “The Internet does not respect collection and national boundaries! Sites in these collections will frequently link to other sites that are not captured as part of a collection, and this can be frustrating for users who inevitably then encounter broken links (Penncock 3).” There are several other issues: legal issues; authenticity issues that deals with quality assurance that affect the preservation and access of the content. One of the biggest technical issues involves the harvesting process because web sites (which is made up of the codes, images, documents, files, and metadata) can only be collected via software.

“Crawler technology has come a long way over the past decade, yet some limitations remain in the types of content crawlers are able to easily capture. Problematic content includes database and dynamically driven content, streamed multimedia files, content accessible only via local site searches where script code is almost impossible for crawlers to analyze, password-protected content, some types of JavaScript-driven menus, and social media (Penncock 11).”

This difficulty partially stems from the constantly evolving web technology. Yet “despite growth in the number of web archiving programs, many institutions still struggle with developing best practices and methodologies to accomplish their goals...in order to address the lack of best practices and to increase awareness of the importance of web archiving as a fundamental to digital preservation, the Archive-It team developed the Web Archiving Life Cycle Model (Archive-It WALCM).” However, what Penncock and the Archive-It don’t mention as part of the limitations of current crawlers and web archiving is the usability and user experience of these software and tools. Even though the digital universe is created and defined by software, a man-made construct, often times programmers have a hard time usefully connecting with users when building services for people “[because they] often have a lot more practice seeing from the computer’s point of view than seeing from another person’s point of view (Sumana).” This means that the service isn’t being delivered because despite the fact that the software *functions*, it is, essentially, not usable.

Usability and the user experience are one of the key factors when programmers are designing their software, yet there is a difference between knowing who is the user and *understanding* who is the user. “Usability refers to the quality of a user’s experience when interacting with products or systems, including websites, software, devices, or applications. Usability is about effectiveness, efficiency, and the overall satisfaction of the user (Usability.gov).” Jakob Nielsen, a well-known consultant from the Nielsen Norman Group, defines usability into having separate quality components: learnability, efficiency, memorability, error frequency and severity, intuitive design, and subjective satisfaction. Determining if the tool is usable is one of the seven qualities Peter Moreville included in his research for creating a good user experience. “User experience (UX) focuses on having a deep understanding of users, what they need, what they value, their abilities, and also their limitations (Usability.gov).” This honeycomb shows that it is necessary to create a unique balance of context, content, and users by *trying* to serve several purposes at once. Any required tradeoff is better made explicitly rather than unconsciously.

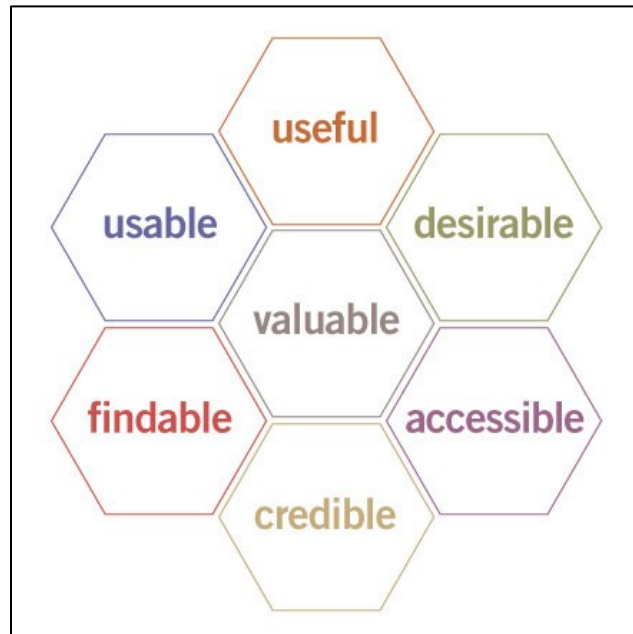


Figure 1: Peter Moreville. “User Experience Honeycomb.”

It may seem like common sense to create something that is usable, but on the web, usability is a necessary condition for survival that many sites and tools lack. “If a website is difficult to use, people leave... If users get lost on the website, they leave. If a website’s information is hard to read or doesn’t answer users’ key questions, they leave... there’s no such thing as a user reading a website manual or otherwise spending much time trying to figure out an interface. There are plenty of other websites available...(Nielsen).”

However, the reality for web archiving is that there usually aren’t other websites available. Just because there is a tool, doesn’t mean that it is useful. In order to ensure that software and tools for web archiving are functional and productive for a user, an analysis comparing a tool, like Archive-It, to Peter Moreville’s “User experience Honeycomb”, is a start to make existing tools more usable for people by determining which sections needs more work and help.

It should be noted that Archive-It is a paid service, which means that expert help from the Archive-It team can be requested throughout anytime of the user experience. However, this analysis will determine how well a user, or ‘partner’,¹ can use the tools provided by the site’s “user-friendly”² design and help documentation. Further more, Archive-It can be considered as an overall tool for web archiving as well as an actual archive. For the purposes of this paper, usability and user experience will be determined based as Archive-It as a tool.

Archive-It is a web archiving subscription service from the Internet Archive that started back in February 2006. It is a web-based application that gives partners the ability to harvest, build, and manage born-digital collections by: providing tools for selection

¹ The user will now be recognized and known as a ‘partner’ according to Archive-It’s terminology.

² This was the description used in their “Learn More” page: <https://archive-it.org/learn-more>

and scoping, including cataloging that uses the Dublin Core metadata format, and storing the digital content on the Internet Archive data centers. Archive-It provides partners with 24/7 access to the tool itself as well as access to the archived content within 24 hours after a capture is complete. This ‘client-side service’ is a common “main acquisition method used because of its simplicity, scalability, and adaptation to client-server environment...this method not only adopts the same position as normal Web users, it also imitates its form of interaction with servers using seed pages, parsing and extracting (Masanes 8).” The underlying software tools are open source and developed by the Internet Archive and provided on IIPC. The Heritrix web crawler captures digital content. Umbra assists with capture, allowing the crawler to access sites in the same way a browser would. Content collected through Archive-It is captured and stored as a WARC file and indexed by Nutchwax (full text search) and Solr (metadata search) software. Two copies of the archived data (primary and back-up) are stored at the San Francisco Data Center and the collections are transferred to the General Archive as a third copy on a hard drive. Partners can view content through the private web application via login/password while the general public and patron can view and access captured content using the Wayback software. Patrons can further the understand their search through post-crawl reviews such as review reports on: seed status, which reports if seed was crawled, shows if the seed redirected to a different URL, or if there was a robots.txt block; seed source, which shows how many documents and which hosts were archived per seed; file type, which lists all the different types of files archived; PDF report that lists all the PDF files found in your crawl; video report that lists all the videos found in your crawl and a QA Report for why specific embedded content on a page may not be captured or viewable and allows for a patch crawl. All archived content display a banner declaring that the current page is an archived web site and time stamped from its capture (wikipedia).

Archive-It definitely presents itself as a “fully hosted solution” since it provides not only selection and acquisition, but also preservation and access to digital content. The Archive-It team tries to diminish the confusion between web archiving and the user experience by providing these tools and help documentation. The User Experience Honeycomb has no particular order or emphasis an importance of one quality over others since affecting one will affect the others. So going in alphabetical order:

ACCESSIBILITY

Accessibility doesn’t necessarily mean whether or not the user has access to the tool. Within user experience, accessibility means whether or not the tool is accessible to people with disabilities. Today it is required for federal services to abide by Section 508, which is the governing principle that requires all government information be accessible to disable users. However, “although accessibility is related to the issues of usability and coding standards, it is possible for a site to be considered usable and standards compliant but NOT accessible and vice-versa (PennState).” It is important to understand that “users have different requirements depending on their situation and capabilities. Catering for these needs and enabling access to the product or system for as many people as possible is the aim of inclusive design.... some factors are: **high stress levels** (short of time, responsibility for children), **low confidence** (unfamiliar with system/tool, first time user), **low perception of safety** (fear of discrimination, compromised data security), **varying abilities** (eyesight, color blindness, motor abilities, hearing impairment), and **language**

barrier (English is not the first language, low literacy, dyslexia) (paznow).” “By making your website accessible, you are ensuring that all of your potentials users, including people with disabilities, have a decent user experience and are able to access you information (Usability.gov).” Case studies show that accessible websites have better search results, reduced maintenance costs, and increased audience reach, among other benefits.

IS ARCHIVE-IT ACCESSIBLE?

While Archive-It has partners in 16 countries, it primarily provides service within the United States. Currently, Archive-It is only in English and Spanish. Both languages are displayed by the login on the upper right corner of the site. The main sections of Archive-It (Home, Collections, Crawls, Reports, Access, Help, Submit a Question) are divided into different tabs so a first time user may be able to navigate from collections to crawls to reports. However, navigating the information within a report does get overwhelming because the seeds are divided and sorted, but it is impossible to sort and search which URLs and information were captured. Right now, the URLs are just listed, which will become stressful to search through when there are 56,879 sites captured with another 75,678 sites queued. And unfortunately, the Help Documentation doesn't provide as much help since it is *less* accessible than Archive-It's site.

Nevertheless, Archive-It is aware of the different types of users and their capabilities and tries to address them in “Archive-It Help”. They developed a K-12 Web Archiving Program, which they update. In a post by Sylvie Rollason-Cass, Archive-It's 5.0 release will focus on “creating an accessible website for all users and have been working toward making Archive-It.org Section 508 and WCAG 2.0 compliant.” This was posted on July 29, 2013 and Archive-It's 5.0 version was just recently released this fall. Archive-It is aware of the flaws in the site and is aiming to fixing them. Ideally, this will also translate and affect the Help Documentation as well.

CREDIBILITY

Credibility is a big issue because authenticity, quality assurance, and integrity are always questioned, especially when dealing with web content and online tools. The Web Credibility Project conducted by Stanford University are beginning to identify design elements and quality components that influence whether users trust and believe content online. As a paid service, the user expects, or hopes to expect, good quality and reliable service from the product and its results.

IS ARCHIVE-IT CREDIBLE?

Archive-It is one of the few subscriptions services actually available for large-scale web archiving. Other services that can be found and used are Archivethe.net from the Internet Memory Foundation and Web Archiving Services (WAS) from the California Digital Library. Archive-It and WAS are client-side services, which is the most widely used acquisition approach for organizations. Client-side archiving is when the web archive is in position of client to gather content from the Web server (Masanes 23). The Web server can generate content from various other servers such as application, database, file servers. Web crawlers like Heritrix or HTTrack act as client and use the HTTP

protocol to gather content responses delivered directly from the server. It is scalable, requires little input from the content owner, and can be highly cost-effective (Penncock 7).^{as} The crawlers come from open source software that is constantly being updated and monitored. Archive-It is a member of IIPC and is part of this maintenance. It seems safe to say that Archive-It is a credible tool and users can trust its service, especially the quality of help as Archive-It hosts informational webinars and in-person and/or email contact. Even though time becomes an issue since webinars are scheduled twice-monthly and emails take up to three days, the Archive-It team does provide answers. And they begin to troubleshoot and research when they can't.

It's always good to see what other users have used Archive-It. Archive-It provides case studies and lists all their partners on their site. One can see that the users range from not only locations, but also scope and missions since Archive-It provides services to universities and colleges, state archives and libraries, museum and art libraries, public libraries and local governments, and national organizations.

DESIRABLE

Desirability not only means if there are users who actually want a tool, but also if the image, identity, brand and other design elements of the tool are used to evoke emotion and appreciation (Usability.gov).

IS ARCHIVE-IT DESIRABLE?

Is Archive-It a tool that users want? The simplest answer to that is yes. It is one of the limited services that helps organizations harvest digital content *and* stores it. Then when analyzing the logo, Archive-It does an efficient job of showing the user and potential clients what it does. It's a tool used for archiving. When one goes on the site, it states that it is "the leading web archiving service for collecting and accessing cultural heritage on the web."



Figure 2: Archive-It logo on the left. Web Archiving Services logo on the right. Archive-It does a more effective promotion of its services compared to the Web Archiving Services. This comparison does not determine or reflect quality of results or services.

FINDABLE

Whether or not a website is findable doesn't just mean if the user can Google it and find it within the search results. Findable refers to if the content of the site is navigable and locatable both onsite and offsite. This usually is linked to the design of the site itself and accessibility.

IS ARCHIVE-IT FINDABLE?

The user must first login to use the tool. Navigating between tabs is doable. However, there are search limitations. Penncock quotes Niu and Stack in that “searching a web archive is different from searching the live web...in the absence of de-duplication strategy, search results from a web archive are often distorted by the presence of multiple copies of identical content harvest in differently crawls. Addressing this distortions poses a significant challenge...Ranking is also different...(Penncock 15)”

A similar perspective should be taken when analyzing the tool. This not only includes the site design, but also the crawls, reports, and help documentation. Finding where the tabs for crawls and help documentation are on the site is easy since they are clearly labeled. Finding specific objects and answers on said crawls, reports, and help documentation isn't as simple. Issues encountered on crawls and reports could be fixed if the help documentation was more navigable.

USABLE

Usability is basically how easy is the tool/software/service to use. As mentioned in the introduction, Jakob Nielsen describes usability's quality components: **learnability**, how easy it is for first-time users to accomplish basic tasks?; **efficiency**, how quickly can an experienced user accomplish tasks?; **memorability**, how easily can a user remember the site in order to reestablish proficiency?; **error frequency and severity**, how often did users make errors?, how serious were the errors?, and how easy was it for users to troubleshoot and recover from errors?; **intuitive design**, does the site have a nearly effortless understanding of the architecture and navigation of the site?; and **subjective satisfaction**, does the user like using the design? Usability encompasses the importance of balancing accessibility, findable content, and credibility by making the tool and help simple rather than simplistic.

IS ARCHIVE-IT USABLE?

As mentioned before, Penncock and the Archive-It team state the issues and difficulties for the web archiving process in general. While Penncock's “Web Archiving: Digital Preservation Coalition Report” and Archive-It's “Web Archiving Life Cycle Model” address these issues and provide suggestions and feedback on how an organization can approach web archiving, neither address whether or not the tool itself is usable for web archiving purposes.

A user *can* use Archive-It. However, that doesn't determine its usability. If the product cannot be delivered to the client and the client cannot get the help needed then the product isn't fully functional. The site itself is mostly usable. A first time user may have to attend an informational webinar to see a live demonstration before understanding and remembering the Archive-It process, which Becky Yoose from Grindell College suggested in the Open Repository 2014 conference as something that should be optional and not required in order to use the tool. But the Archive-It team does provide help via “Submit a Question” and these webinars to make the use of this tool efficient. The user can determine the frequency of crawls and receive reports 24 hours later once the crawl is completed. The steps could be clearer (on both the site *and* the help documentation), but ultimately the user is satisfied by the end product of these crawls even if the user

experience can be a bit frustrating, especially when trying to fix and troubleshoot any errors. Its usability can be greatly improved if the help documentation was more accessible, findable, and overall usable. Ms. Yoose had stated that a common issue affecting user experience is documentation. Documentation on web archiving is growing, but documentation on *how to use* tools, software, and the process is still very limited: “documentation, if you can find it in the first place, is vague, outdated, or goes over the heads of the staff reading it (Becky Yoose).” Help documentation is part of the user experience and a primary factor for the usability of the tool. “Archive-It Help” does get updated, but important sections such as the “Archive-It Guide”, which gives step-by-step instructions on how to use Archive-It hasn’t been updated since May 9, 2013, despite the fact that Archive-It is already on its 5.0 version release. And while Umbra has made archiving social media more doable, there still needs to be some troubleshooting and research done for when dealing with social media sites like Instagram in order to also get the comments and tags that are included with the photo or at least to understand why.

USEFUL

Usefulness is the result of when the utility provides features the user needs and those features are easy and pleasant to use. It is determined by whether or not the tool fulfills a need.

IS ARCHIVE-IT USEFUL?

Despite the fact that the help documentation could be better, Archive-It fulfills this need in large-scale web archiving by providing tools for acquisition, selection, and storage. Yet the inability to sort and analyze the crawls and reports in an advanced search, search through the digital content, or acknowledge the errors on the help documentation, limits the usefulness of the tool. I would argue that a better understanding and search for the digital content is needed. While the visual product of the archived web page is desired as an end result, if the user cannot understand, search, or sort the digital content that comprises the site then the user is just getting a bunch of data that they can’t read or understand and will have to sort out in the future.

VALUEBALE

Gauging the value is definitely a subjective question. “Is the end product and experience valuable to users?” will vary by user, experience, and the actual use of the tool. However, regardless of the type of content captured, web archiving and these organizations and services that focus on web archiving are valuable because it acknowledges and archives the online human record. The value is determined by “Why even do web archiving?” explained in the beginning of this paper: “...the web archiving question should be viewed in the perspective of cultural heritage preservation... By nature, Web content is ephemeral: sites disappear continually and are frequently updated, involving the disappearance of what often is very valuable online information and Web content (IFC).”

IS ARCHIVE-IT VALUABLE?

The purpose of Archive-It's service is definitely valuable. It provides a necessary service within a growing field in archiving. It facilitates web archiving with acquisition tools (maintaining and monitoring the Heritrix web crawler and Umbra); suggests possibilities and guides in the selection and scoping process; is able to host and store high quantities of digital information captured and provide access to not only the client, but also the public.

Nevertheless, its value as a tool and the value of the content can increase if action was taken to target the flaws and errors in order to create a better user experience.

Conclusion: What can be done?

What makes a good and bad experience? The User Experience Honeycomb and breakdown of usability provides a good sense of what can create a good experience. Unfortunately, the answer is ultimately subjective depending on the user and situation.

Now there are several usability testing and laboratories to improve user experience. Usability.gov is one of the leading resources for user experience best practices and guidelines. It takes into account the business goals and user goals by developing these best practices in order to promote improving the quality of the user's interaction with and perceptions of the product and any related services. User involvement can lead to the most successful results when designing the product through direct feedback, user testing, observation or informed evaluation using previous gathered information. As Becky Yoose stated in her presentation, "A Tale of Two Communities," vendors need to: "...test their software in places that don't look like theirs. If [users] can't implement the stack, it must be fixed *on their end*. The vendor have a responsibility to the community to make its software functional as the users in the community have a responsibility to send bug reports and to help troubleshoot." The vendor, for example Archive-It, must be aware of not only the presentation of the content, but also the user's access to that content: "...the presentation of content is at the viewer's discretion and is further shaped by the viewing tools used (browsers and other applications). In other words, though the content may be the same Web page, two viewers using different browsers...may ultimately see different pages...(Jones 55)."

In the case of Archive-It, there is a definite user involvement when updating the crawlers and web archiving capabilities. It was through feedback that Archive-It began troubleshooting how to archive a Tweet because it was becoming a record for libraries and archives. Archive-It definitely provides an overall good user experience with service and product value. Yet as a result, any bad experience is highlighted because of the jarring difference from the generally smooth process. What can be done to remediate these bad user experiences is to address the weak points in design, quality assurance, and documentation. Throughout this analysis, it seems that the most direct and practical improvement that both Archive-It and the user can do is improving the help documentation.

Archive-It needs to update its help documentation, especially to reflect the latest releases from the past year. Even though crowdsourcing information and users' comments runs a risk of lessening the credibility and quality of the guidelines, these are good approaches to usefully connecting with the user by having other users relate to each other with their own problems and experiences. Another possibility would be to illustrate

the potential errors a user will encounter to prepare them during their process. BAVC's AV Artifact Atlas is a great example of how a community-based, online resource helps users identify and diagnosis artifacts and errors in analog to digital practices. For example, it might be beneficial to show the user how an archived Instagram page would look like since currently Umbra cannot capture comments or enlarged images for Instagram pages:

Instagram

Our 4.9 Release of Umbra has improved our ability to capture and replay Instagram pages. There are currently no specific crawl modifications necessary for Instagram crawls. We are now able to replay two scrolls of the dynamically loading content for Instagram pages.

You may not be able to enlarge images or see the comments and likes for an image.

Our engineers are now working towards further improving our capabilities for crawling Instagram. Please let us know if you encounter any issues or have questions regarding crawling Instagram.³

Concise and clear documentation should be done properly and this is usually a difficult task when dealing with technical issues to a mostly non-technical user. And this is a problem that other services, not just Archive-It, experience. In a recent blog post, a similar situation was mentioned about GitHub: "I think a lot of people are intuitively intimidated by using git because it can be hard to understand or because they don't know how to read code. But not being able to read code is okay because you can still make valuable contributions in English! One of the biggest sore spots for most open source software (and even proprietary software) is a lack of clear, concise documentation (Ashley Blewer)."

Concise and clear documentation is an issue. Web archiving is an issue in general too. And Archive-It helps address this issue. With proper documentation and collaborative help, a proactive approach can begin with web archiving by giving interested users these open source tools provided by IIPC in order to capture this vast amount of digital content available on the web. First there should be proper documentation available: "don't just do it, do it right (Becky Yoose)." Write the Docs (docs.writethedocs.org) is an online resource to help develop clearer and concise documentation for, mostly technical, processes. Knowing that the IIPC and Digital Preservation Coalition (DPC) provide resources and articles or the GitHub and forums are communities where questions can be asked and answer, can also help direct users.

Ironically, in order to search and find everything desired to archive the web, the user needs the Web, especially for access purposes. For example, on December 11, 2014, the public user was unable to access the digital content captured, stored, and hosted by the Internet Archive, the WayBack Machine, and Archive-It.

³ This latest post was updated December 10, 2014, a week after their first informational webinar with users.

“The big storm in SF has knocked out power to our main data center, so the site will be down for a while. We’ll keep you posted here! (@internetarchive).”

Failure to find the user’s desired results via Google, YouTube, or some sort of forum can (and often times will) result to a bad user experience. Failure to access desired results could lead to a bad user experience because of this denial of immediate gratification. It is impossible to guarantee a good user experience every time a user uses the Internet. Similarly, services and tools like Archive-It try to satisfy the user with a good user experience with desired results, but neither Archive-It nor Internet Archive can be expected to harvest and store all of the digital content found on the web.

Acknowledgement of web archiving as digital preservation within organizations and individuals needs to be established. Our society expects so much from the web and in reality Masanes suggest a possible idea for the future of web archiving:

“...Another possible evolution is the generalization of this to enable every Web user to participate actively if she or he wants, in the take of archiving the Web. The main incentive for users is, in this case, to organize their own personal Web memory to be able to refer back later to stable content, but also to mine it and organize it as a way to fight the “lost in cyberspace” syndrome (6).”

An evolution I feel that can be done with proper documentation and collaborative help in web usability.

BIBLIOGRAPHY

International Internet Preservation Consortium. "Why archive the web?"
netpreserve.org/web-archiving/overview

International Internet Preservation Consortium. "Why archive the web?"
netpreserve.org/web-archiving/videos/why-archive-web

Internet Memory Foundation. "Web archiving: What is web archiving?" *About*
http://internetmemory.org/en/index.php/about/web_archiving

Masanés, Julian. "Web archiving: Issues and methods." *Web Archiving*. New York: Springer, 2006.

Pennock, Maureen. "Web-Archiving." *DPC Technology Watch Report*. Digital Preservation Coalition. 13-01 March 2013

Jones, Steve and Camille Johnson. "Web Use and Web Studies." *Web Archiving*. New York: Springer, 2006.

Nielsen, Jakob. "Usability 101: Introduction to Usability." *Nielsen Norman Group*. 4 January 2012, <http://www.nngroup.com/articles/usability-101-introduction-to-usability/>

Becky Yoose. "A tale of two communities." *Open Repositories 2014*

Sumana Harihareswara. "UX is a social justice issue." *Code4lib* 25 March 2014
http://wiki.code4lib.org/index.php/2014_Keynote_by_Sumana_Harihareswara

Internet Archive (internetarchive). "The big storm in SF has knocked out power to our main data center, so the site will be down for a while. We'll keep you posted here!" 11 Dec. 2014, 7:59 a.m. Tweet.

Usability.gov. "User Experience Basics." *What & Why of Usability*
<http://www.usability.gov/what-and-why/user-experience.html>

Usability.gov. "Accessibility Basics." *What & Why of Usability*
<http://www.usability.gov/what-and-why/accessibility.html>

Usability.gov. "Usability Evaluation Basics." *What & Why of Usability*
<http://www.usability.gov/what-and-why/usability-evaluation.html>

Rollason-Cass, Sylvie. "Archive-It.org Individuals with disabilities accessibility updates." *Archive-It Help*. 29 July 2013.

<https://webarchive.jira.com/wiki/display/ARIH/Archive-It.org+Individuals+with+Disabilities+Accessibility+Updates>

PennState. "Standards: Accessibility vs. Usability vs. W3C Web." *The Penn State University*. 2014. <http://accessibility.psu.edu/usability>

"Internet Archive." Wikipedia: The Free Encyclopedia. Wikimedia Foundation, Inc. Last modified 5 Dec. 2014. http://en.wikipedia.org/wiki/Internet_Archive

Paznow. "User centered design." *Multi-disciplinary design*. 2010
<http://www.paznow.com/ucd/>

The Archive-It Team. "The Web Archiving Life Cycle Model." *Internet Archive*. March 2013

Blewer, Ashely. "Non-technical Person's Guide to Becoming an Open Source Software Contributor via GitHub." *GitHub*. 3 November 2014.
<http://ablwr.github.io/blog/2014/11/03/non-technical-persons-guide-to-becoming-an-open-source-software-contributor-via-github/>

Pictures

Moreville, Peter. *User Experience Honeycomb*. Digital image. *Semantic Studio User Experience Design*. Semanticstudio.com, 21 June 2004. Web. 7 Dec. 2014.
<http://semanticstudios.com/user_experience_design/>.

Internet Archive. *Archive-It*. Digital image. *Archive-It*. Archive-it, Feb. 2006. Web. 7 Dec. 2014. <archive-it.org>.

California Digital Library. *Web Archiving Services*. Digital image. *Web Archiving Service*. The California Digital Library, n.d. Web. 7 Dec. 2014.
<<http://webarchives.cdlib.org>>.