Submitted on: 01.06.2017

News Media Satellite Meeting
April 26-28, 2017
National and University Library of Iceland, Reykjavik

TRACKING USER EXPERIENCE OF DIGITAL NEWSPAPERS ON THE PORTAL TO TEXAS HISTORY

Ana Krahmer, Ph.D.
Director, Digital Newspaper Unit, University of North Texas Libraries
ana.krahmer@unt.edu

Abstract:

A widespread debate in digital newspaper access and preservation circles is how users work with newspapers and whether they prefer article-level segmentation or full-page views of news content. While professionals involved in making digital newspapers accessible constantly discuss this issue, little attention has been paid in actually observing users working with newspapers. This paper will discuss the research findings of a user experience project directly related to how people interact with newspapers on The Portal to Texas History. These findings include eye-tracking, verbal, and performance data that will help fill in information about what people do when they use digital newspapers in the web environment. The Portal to Texas History displays search results in newspapers as full-pages, rather than with segmented article review. This paper will represent the first of two research projects comparing user experience with article-level segmentation versus full-page display.

Keywords: digital newspapers, preservation, user experience, usability

Introduction

Created in 2002, the Portal to Texas History out of University of North Texas Libraries (UNT) hosts over 900,000 unique objects related to Texas history. The Portal serves as both a preservation and access infrastructure for all content that it hosts. This content is accessible free of charge, with no login access required to work with the materials. Since 2014, UNT has been a service hub for the Digital Public Library of America (DPLA). UNT Libraries, through its Digital Newspaper Unit, has been

selected for four awards cycles to participate in the National Digital Newspaper Program (NDNP), sponsored by the National Endowment for the Humanities (NEH). At the time of this writing, the Texas Digital Newspaper Program collection hosts 1,117 newspaper titles, representing 222 Texas counties, comprising 5.1 million newspaper pages.

In Summer 2016, the Digital Libraries' Division at UNT released a new interface to The Portal to Texas History. With this release, Digital Libraries' staff with backgrounds in user experience research initiated a project to assess user experience with the new Portal interface. The research team is comprised of Ana Krahmer, Dianne Jansing, Jacob Mangum, and Robert Lay. I, Ana Krahmer, led this team in part because my own doctoral research studied The Portal to Texas History, but even more because I wanted the opportunity to observe users working with newspapers on the Portal. Since UNT first began to make newspapers available on the Portal, we have hosted full-page views, with the option to zoom into newspaper page areas and create permanent links to those areas, as if to display just a clipping. Over the past six years, I have attended multiple conferences that dealt with building access to newspaper collections, including the National Endowment for the Humanities' National Digital Newspaper Program (NDNP) annual meetings, "Beyond NDNP" work sessions, and various genealogical conferences, and one topic of conversation in which these meetings invested a significant amount of time was article-level segmentation for displaying newspapers versus full-page newspaper views. If these meetings have taught me anything, I have learned that people who work in newspaper digitization are devoted to article-level segmentation when their libraries use it, but those whose libraries do not are cautious about committing to this type of display because article-level segmentation can incur additional charges and can take longer to prepare. This paper explores the beginning steps of a user experience test on The Portal to Texas History which examined usage by participants familiar with both the full-page interface of the Portal as well as those interfaces of other genealogy websites that provide article-level newspaper display.

Review of Literature

Extant digital newspaper collections in the U.S. emphasize access and accessibility as a key feature in their mission, whether those collections are hosted by libraries or corporations. University of North Texas' Texas Digital Newspaper Program (TDNP) has a stated mission of, "The Texas Digital Newspaper Program (TDNP) partners with communities, publishers, and institutions to promote standards-based digitization of Texas newspapers and to make them freely accessible via The Portal to Texas History" (TDNP, 2017). Chronicling America, the U.S. digital newspaper repository sponsored by the National Endowment for the Humanities, states something very similar to the TDNP mission: "is a long-term effort to develop an Internet-based, searchable database of U.S. newspapers with descriptive information and select digitization of historic pages" (Chronicling America, 2017). The Utah Digital Newspapers represents a collaborative effort between Brigham Young University, Utah State University, and Salt Lake Community College to support distributed statewide aggregation and hosting for "Creating Citizen Historians" (Utah Digital Newspapers, 2017). Herbert and Estlund highlight in their 2008 article, "Bringing the Past to the Present," that the Utah Digital Newspaper practice of article-level segmentation services a few different, significant purposes: "Segmenting pages into separate articles has several advantages: Articles are presented in search results, article images are presented for viewing, and OCR is improved because there are more consistent fonts within an article and hyphenated words are more easily conjoined. We find these reasons compelling enough to justify the additional expense of segmentation. Indeed, 67% of UDN users rate the search accuracy as "good" or "excellent." (p. 35). Reakes and Ochoa, out of University of Florida, in 2009 state, "Preliminary investigation done by the researchers indicates that, in the case of non-commercial digital newspaper collections, little emphasis has been put on the usability and end user satisfaction associated with the final products of these efforts" (p. 1). In the U.S., even now, large-scale digital newspaper programs have been in an apparent race to build access to newspapers.

Those U.S. institutions involved in the newspaper preservation community stress numbers of pages and program size relative to other institutions, with core digitization and preservation vocabulary that emphasizes competition: who has the most pages, the largest number of OCR cores, the most terabytes of data. However, even as of the time of this writing--an embarrassing eight years after Reakes and Ochoa's admonition--little self-evaluation analysis of user experience has been made publicly available, representing a dangerous void in cross-institution newspaper access practice. The University of Florida work with their digital newspaper program should have represented a foundation of UX work on which other institutions could build, and perhaps it did represent this, but we in the newspaper preservation and access realm have done a poor job publishing our research when it does exist. UX self-evaluation--and discussion of evaluation results--should be a natural next step in building access to digital newspaper collections, now that the process of digitizing newspapers and making them available is entrenched, through the work of such massive news preservation undertakings as Chronicling America, the Utah Digital Newspapers, the University of Florida Digital Newspapers, and University of North Texas' TDNP.

2016-2017 PORTAL TO TEXAS HISTORY USER EXPERIENCE RESEARCH PROJECT

One acronym summarizes everything usability and UX research look for: MEELS, or Memorability of a given product, Efficiency of a given product, Errors in a given product, Learnability of a given product, and Satisfaction with a given product (Still, 2011, p. 6). While usability addresses issues of the ability of a product to fulfill usage needs of its consumers, user experience is an emphasis on user needs and satisfaction (Bevan, 2009, p. 1) in working with a given product. The 2016-2017 Portal to Texas History study centers on actual user experience, gathering data on what users see, what they say, and what they do. Eye-tracking data provides visual information about what users see as they work with The Portal to Texas History and with the newspapers, specifically. Task recording with Morae TechSmith provides observational data. End-of-task retrospective recall and active intervention questions explore what users say about newspapers on the Portal. This triangulation of data ensures accuracy and rigor within the test design. This study has been approved by the University of North Texas Institutional Review Board. The study will continue until at least December 2017, so the plan and results discussed here represent a beginning of what will be at least a year-long study.

The first four tasks center on searching for and inside newspapers and browsing around newspaper issues and collections (see Table 1). The testing team deliberately planned these tasks to build user confidence because, if tasks begin at too high a level, users will walk away without completing everything.

Task Order	Task Text	Desired Outcome
Task 1	From the Google page, please perform a search for "Abilene Daily Reporter," and visit the first through third search results.	• First task should be easier, to boost users' confidence • To examine how users move between a mainstream search engine directly into an issue of a newspaper on the Portal

		• This task also will load Chronicling America should the users visit that
Active Intervention Question, Task 1: When you moved from the Google page and visited the search results, can you explain what kind of websites you thought you were visiting?		
Task 2	Please visit the link that goes from Google to The Portal to Texas History. Search inside this issue for "Smith." Take some time to look around until you feel you understand what is going on in the page.	• To observe how users orient themselves from a search engine to a single newspaper issue. • To observe what users choose to examine first inside a single issue.
Active Intervention Question 1, Task 2: What was your thought process as you started to orient yourself on the page? Retrospective Recall Question 1, Task 2 (asked if users moved around inside the newspaper page or through the issue): What were you looking for as you zoomed in and/or flipped through the pages?		
Task 3	Staying on this website, click wherever you want to go next from this page. Explore the website.	• To observe where users will visit next • To see, via the eye-tracking data, where users look from one newspaper page to explore somewhere else
Active Intervention Question 1, Task 3: What made you choose where you explored on the website? Active intervention Question 2, Task 3: What were your first thoughts as you explored on the website?		
Task 4	Click on the next tab, and perform an advanced search. Use the given prompt document [provided to user] to reference which search terms you'll enter into the search fields.	• To examine how users interact with newspapers when performing a very specific search • To observe how Advanced Search fields available on the Portal support user success • The most difficult task, with the first three tasks planned to build user confidence
Active Intervention Question 1, Task 4: Can you explain how Advanced Search is different from other search methods you've seen in the past? Active Intervention Question 2, Task 4: Can you explain to me what each of these [Advanced Search] fields mean?		

Task 5	From the home page in the next tab, locate the "Featured Partner" and tell us who it is.	• Final task should help users leave with confidence
Active Intervention Question 1, Task 5: What made "Featured Partner" stand out (or not stand out) to you?		

The Discovery Process in designing any kind of UX test serves to gather information about how real-life users work with the product you plan to test. Discovery, according to Brian Still, functions for the usability research in a similar way to ethnographic research: quietly observing users working in day-to-day situations with materials, gathering data on use cases, and learning what users think as they perform these daily routines (Still, 2011, p. 50). For the Discovery Process of this test instance, Jansing and Krahmer, the two primary researchers on the project, sat down with two users unfamiliar with The Portal to Texas History and asked them to perform research for one hour. The users could navigate anywhere they wanted, but they had to pose a research question and look for results to answer that question. We created tasks 1-5 directly from the results of these Discovery sessions.

Populations Tested

Sampling was purposive for this test instance, with participants who either have worked with The Portal to Texas History before or who represent the target audience of the Portal and TDNP, such as genealogists, historians, educators, or students. We recruited this first round of participants over The Portal to Texas History listserv. All personally-identifying information about users is stripped in this data, and users are coded to ensure anonymization.

Preliminary Results

Because these results represent the first round of testing and since I'm discussing results from just a few of the tasks, I will summarize briefly trends users seem to be demonstrating.

The question I deliberately pose when a user starts looking up and down a newspaper page or browsing through a newspaper issue is a Retrospective Recall question. In usability, Retrospective Recall questioning is a methodology that directly interrogates a user about unique choices that he or she made during testing (Still, 2011, p. 61). Retrospective Recall target specific behaviors of the individual users. I could only ask this kind of question when a user actually navigates the page and the issue.

For task 2, when users start searching inside an issue of *The Abilene Daily Reporter* for the name "Smith," six results will appear on six different pages, and users can browse through the pages to find out what they say, if they so choose. When users start reading through multiple pages, I ask them, "What were you looking for as you zoomed in and/or flipped through the pages?" In responding, one user said, "At first, I was trying to get more information about this first 'Smith' result, but then I wanted to read the rest of the newspaper page to center myself on when it was published and what world events were happening that the Abilene newspaper was writing about for this specific issue." Another user reported, "I like having the context for what I'm trying to locate. It's nice to know the background for that newspaper issue." In terms of performance, these users clearly perused the full newspaper page, as is evident in figures 1-2, and in one case different pages in the same issue, with the page zoomed-in to show the text. As this is the first of what will be three to four additional rounds of testing, this data is not complete, and I share it as a preview of what this research project will find in the future.

Figure 1: Eye-tracking data of user examining a newspaper page for "what world events were happening . . ."

Figure 2: Eye-tracking data of another user examining the "context for what I'm trying to locate."

Figures 1 and 2 demonstrate concentrated viewing, with the brighter areas of the heatmap representing the longer amount of time a user spent fixating on an area of the page. A fixation is defined as the amount of time an eye stays looking in one area. The fixation time is set to 200ms with the data window set on 2000ms, and the fixation time-length is represented on a color scale ranging from light blue to yellow to orange to red. The heatmaps represent actual reading, with concentrations of attention for these two users varying according to what drew their attention on the page.

CONCLUSION & RECOMMENDATIONS

Because the data-gathering process on this project is in its very earliest stages, there is no observation coding to report on at present. Over the next year, the testing team will try to complete full-scale testing on 30 total users. Should this be successful, the entire team will evaluate and report on the full results in a much longer and more cohesive document. This paper is simply a brief report on how this test instantiation looks at one aspect of newspapers in the TDNP.

My own interest as a UX researcher stems directly from my position as the Digital Newspaper Program Director out of UNT Libraries. In this position, I work with patrons on a daily basis, and many of these patrons represent newspaper collection contributors whose efforts have made entire runs of newspapers accessible via The Portal to Texas History. In addition, I work with representatives from public and private funding agencies, as well as publishers, who believe that access to newspapers is critical to preserving community identity, and these are also groups who value positive user experience and research into how TDNP can best support this value.

For anyone interested in initiating their own UX research on their own state-wide digital newspaper repositories, my final recommendations are as follows:

1) Set useful goals: UX test planning begins with goals that address actual issues or questions. For newspaper repositories, goals should focus on how users actually perform research and locate information. For example, one of our goals for the newspaper portion of the project was to understand whether users were actually manipulating and utilizing objects on the Portal in the way in which they were intended to be viewed. The newspapers on the Portal contain multiple use features to help improve people's experiences, and understanding if and how they use them is a significant first step in knowing how our newspapers serve the public.
2) Discovery combined with goals will determine the direction: Digital newspaper collections are used by so many different types of people, and these are people who frequently use every type of digital newspaper repository they can access. Discovery will help you narrow down how you want to evaluate your repository, whether you want to tweak your goals, and the combination of Discovery and goals will develop your tasks.
3) The full-scale test is not always the best choice: User experience work comes in many different flavors, including card-sorting, paper-prototyping, paired cognitive walkthrough, A/B testing, among them. The Portal redesign testing combines concepts from multiple UX methodologies to ensure research triangulation, and because we had access to high-fidelity equipment and were able to set aside the time to do it, this research team decided to perform the large-scale test.
4) Even a little bit of testing is good: Something many, if not all, UX professionals say that is completely true, is that a little bit of UX research data can go a very long way toward helping you target areas for improving your design processes and access interfaces. The payoff for a little bit of UX work is high.
5) Share your results: If we want to learn more about how people function in the digital newspaper preservation community, please share your work. We would all love to learn from your experiences!

Acknowledgments

The author wishes to thank the research team who has worked so hard on The Portal to Texas History Redesign UX Project: Dianne Jansing, Jacob Mangum, and Robert Lay. Every member of this group contributes important talents and knowledge, and I am proud to have worked with each of these people.

REFERENCES

Bevan, N. (2009). What is the difference between the purpose of usability and user experience evaluation methods? UXEM'09 Workshop, *INTERACT 2009*, Uppsala, Sweden.

Herbert, J. & K. Estlund, (2008). "Bringing the Past to the Present." *Online Magazine.* 32 (4), 32-37. http://hdl.handle.net/1794/9914

Library of Congress. National digital newspaper program. *Chronicling America* (2010). Accessed April 15, 2017. Retrieved from http://chroniclingamerica.loc.gov

Reakes, P. & Ochoa M. (2009). Non-commercial digital newspaper libraries: Considering usability. *Internet Reference Services Quarterly*, 14 (33-34), 92-113. http://dx.doi.org/10.1080/10875300903336357

Still, B. (2009). *English 5388: Usability workshop handbook*. Lubbock, TX: Texas Tech University.

University of North Texas Libraries. (2017). Texas digital newspaper program. Accessed April 16, 2017. Retrieved from https://texashistory.unt.edu/explore/collections/tdnp

Utah Digital Newspapers. (2017). About Utah digital newspapers. Accessed April 16, 2017. Retrieved from https://digitalnewspapers.org/about