User-Centered Collaboration for Archival Discovery

Session 403 / #s403
SAA 2017
Portland, Oregon

User-Centered Collaboration for Archival Discovery

10:00 - 10:30 Presentations

- Archival Discovery at NYU
- ArchivesSpace Public User Interface
- ArcLight

10:30 - 11:00 Facilitated group discussion

Presenters and Facilitators

Chair: Mark Matienzo, *Collaboration & Interoperability Architect, Stanford University Libraries*

James Bullen, *Founder, Hudson Molonglo*

Wendy Hagenmaier, *Digital Collections Archivist, Georgia Institute of Technology*

Emilie Hardman, *Research, Instruction, and Digital Initiatives Librarian, Harvard University*

Susan Pyzynski, *Associate Librarian of Houghton Library for Technical Services, Harvard University*

Mike Shallcross, *Assistant Director for Curation, University of Michigan Bentley Historical Library*

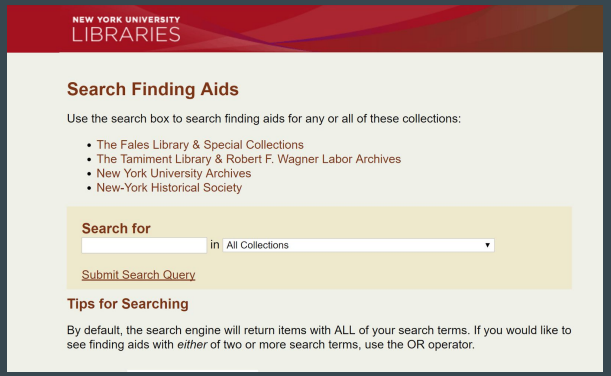
Sally Vermaaten, *Manager, Archive Solutions, Gates Archives*

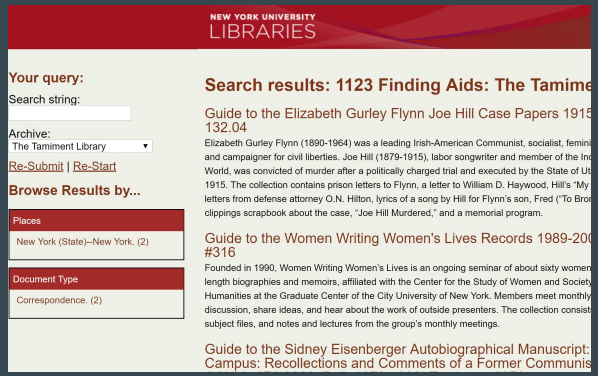
Archival Discovery at NYU

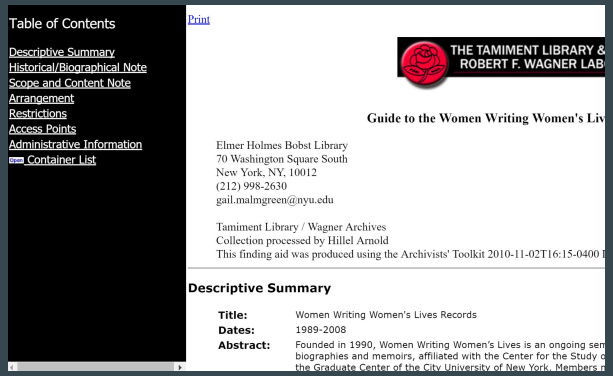
Finding Aids ‘Bridge’

Sally Vermaaten
Gates Archive

2013







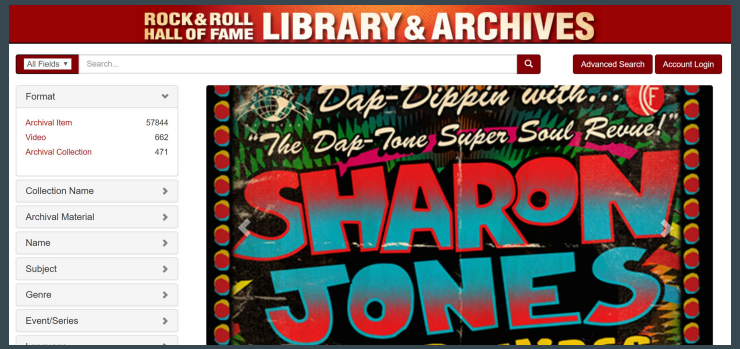
specialcollections.library.nyu.edu | https://github.com/NYULibraries/findingaids | Sally Vermaaten

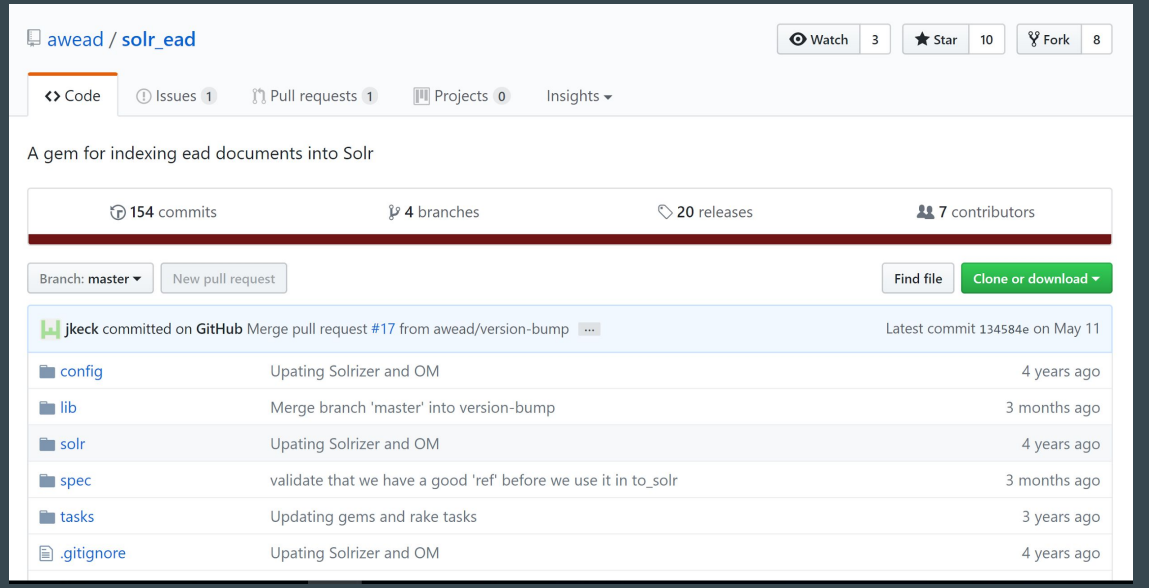
2013 - Methods

- Literature review
- Ideation workshop and affinity grouping
- Peer system assessment
- High-level requirements
- Personas (started)

Proof of concept

blacklight

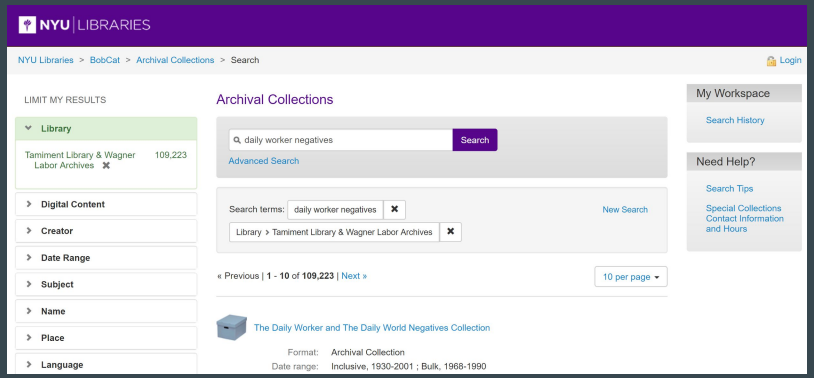


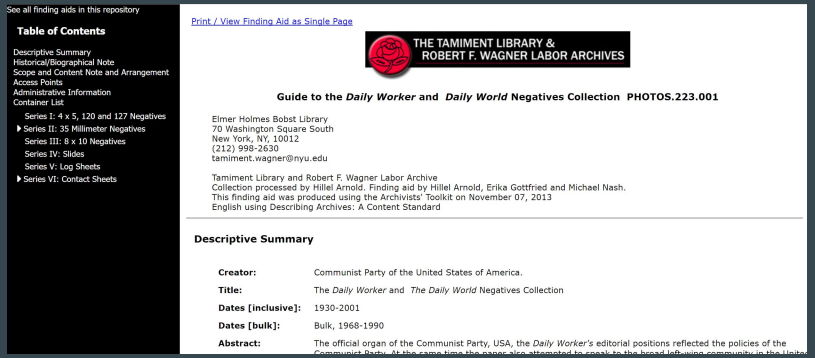


Prioritizing discovery with other system needs

1. Archival search and browse - Blacklight
2. Collection management - ArchivesSpace
3. Request and workflow management - Aeon
4. Special collections discovery - ???

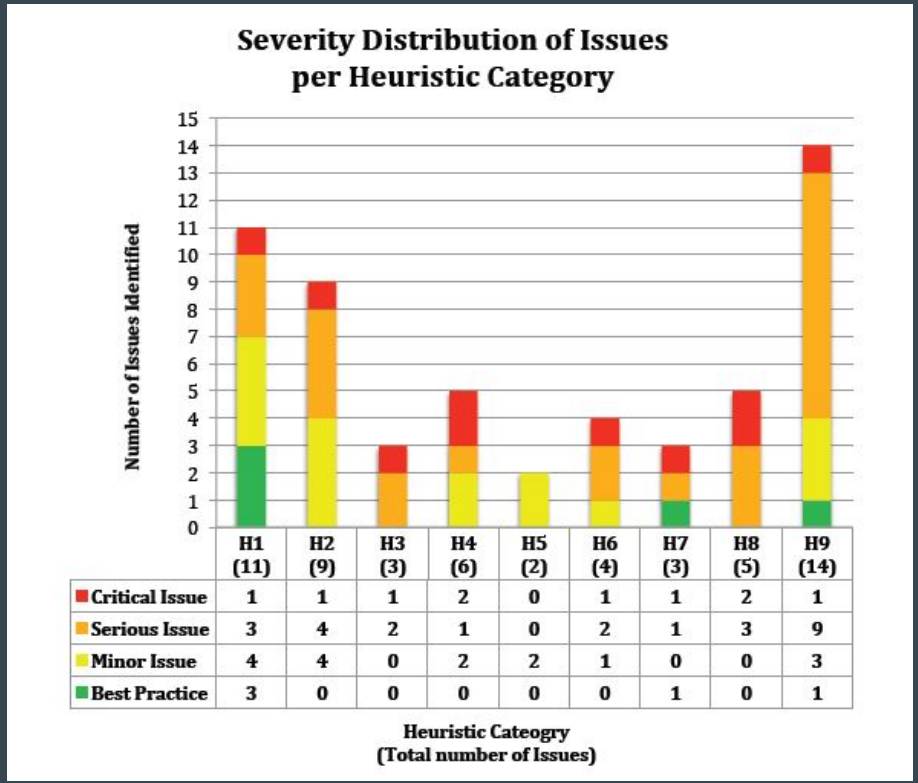
Bridge to a better user experience



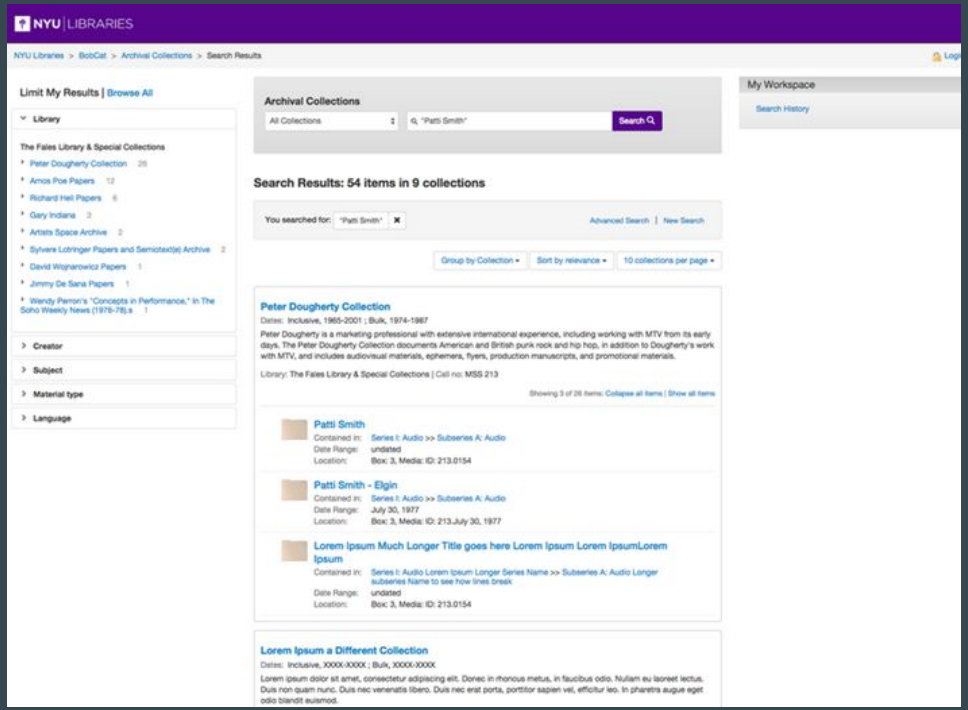


Methods - Heuristic evaluation

H2	**Language** The interface should speak the users' language, with words, phrases, and concepts familiar to the user rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.	Q1: Is the language used on the site easy to understand?
		Q2: Is the terminology consistent, both within the Archival Collections Finding Aid and compared with BobCat?
H3	**User Control** Users often choose interface functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Supports undo and redo.	Q1: When the user makes errors, is it easy to exit the unwanted state without going through unnecessary dialogue?
H4	**Cognitive Load** The user should not have to remember information from one part of the dialogue to another. Instructions for the use of the system should be viselike or easily retrievable whenever appropriate.	Q1: Are actions, options and objects always visible without the user having to remember them? (e.g. list of previous searches).
H5	**Flexibility & Efficiency** Accelerators improve efficiency for interactions performed frequently (e.g. redo a previous search).	Q1: Are accelerators used on the site?
H6	**Aesthetic** Dialogues should not contain information that is irrelevant or rarely needed.	Q1: Is irrelevant or unnecessary information minimized?
H7	**Errors** Dialogues should not contain information that is irrelevant or rarely needed.	Q1: Are errors prevented by confirmation options?
		Q2: Are error messages constructive and in plain language?
H8	**Help** Error messages should be expressed in plain language (no codes), indicate the problem precisely, and suggest a solution constructively.	Q1: Is help documentation easy to find and focused on the task at hand?
		Q2: Does it contain concrete steps, and is it not too large?



Methods - Wireframing and feedback



Prioritizing discovery with other system needs

1. Archival search and browse ✓
2. Collection management - ArchivesSpace
3. Request and workflow management - Aeon
4. Special collections discovery - ???

2017 - Special Collections Discovery System

Takeaway #1

Where possible, share your user assessment data, system evaluations, design artefacts, as well as code. Seek to build on the work of others rather than re-inventing wheels.

Takeaway #2

Identifying and documenting user needs can be both an input into and impetus for software development and improvement projects.

Takeaway #3

Consider proof-of-concept approach and incremental, phased roll outs of different components.

ArchivesSpace Public User Interface User Requirements and Assessment

susan pyzynski

associate librarian for technical services
houghton library, harvard university

emilie hardman

research, instruction & digital initiatives librarian
houghton library, harvard university

HARVARD LIBRARY

Beginning the Design Process, 2015

- ArchivesSpace starts the process to create a new PUI
- Mark Custer (Yale) agrees to lead the membership driven process
- May: a first call for volunteers goes out, a 14 member working group is formed
- Cherry Hill chosen as the design firm
- June: the PUI Enhancement Working Group has first virtual meeting, design process runs through December 2015

HARVARD LIBRARY

ArchivesSpace PUI Design Working Group

Mark Custer, Yale University, Chair

Susan C. Pyzynski, Harvard University

Linda Hocking, Litchfield Historical Society

Susan Luftschein, Univ. of Southern California

Maura Carbone, Brandeis University

Matt Francis, Penn State University

Mariella Soprano, Calltech

Dara Flinn, Rice University

Cate Putrirskis, Ohio State University

Scott Schwartz, University of Illinois

Cory Nimer, Brigham Young University

Krista Ferrante, MITRE Corporation

Claryn Spies, Yale University

Elisa Piccio, Caltech

Jessica Dowd Crouch, Univ. of South Carolina

HARVARD LIBRARY

Collaboration

- Virtual meetings every two weeks with the working group and Cherry Hill
- Used GoToMeeting and Slack for the collaborative design process
- Careful documentation of design and decision-making through a public wiki https://archivesspace.atlassian.net/wiki/spaces/ADC/pages/66355216/Design+Phase
- Sought outside comments and contributions
- Reviewed and integrated all past functionality requests from members

HARVARD LIBRARY

First Round of User Testing

- Determined we should incorporate user testing of original ArchivesSpace front end in design process
- Harvard, Rice, and Yale performed user testing in conjunction with Cherry Hill
- Results were incorporated in the final Cherry Hill design
- Decided to do continuous user testing during the development phase of the PUI

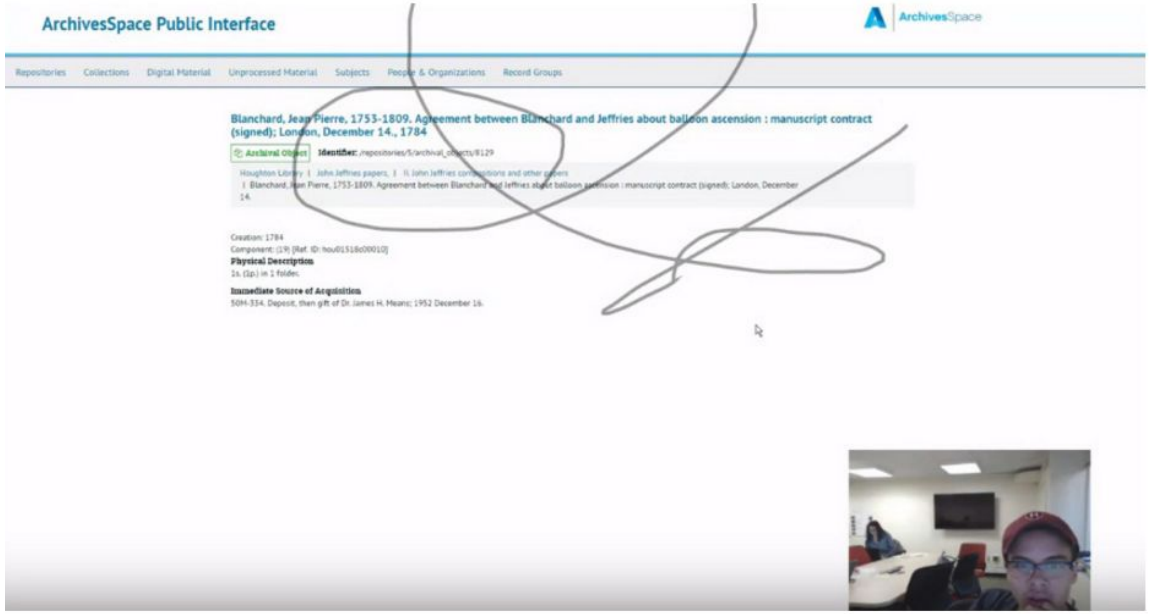
HARVARD LIBRARY

User Testing: PUI Development Phase

- Development phase runs January 2016-June 2017
- Smaller Development Group formed: members from ASpace/Lyrasis, Hudson Molongo, Yale, Harvard
- Harvard volunteers to take the lead in user testing
- User testing carried out: July-November 2016
- Incorporation of test results into development process and decisions

HARVARD LIBRARY

> "From the moment I saw the site I was automatically like I'm going to be frustrated and, here, exactly, it's just not giving me the right details."



screengrab illustrating mouse trail of the participant expressing frustration

HARVARD LIBRARY

UX at Harvard

User Research Center (2 FTE, including Senior Consultant)

Trained User Experience Consultants embedded in 7 functional areas through Library

HARVARD LIBRARY

Methods

Moderated task-based usability testing

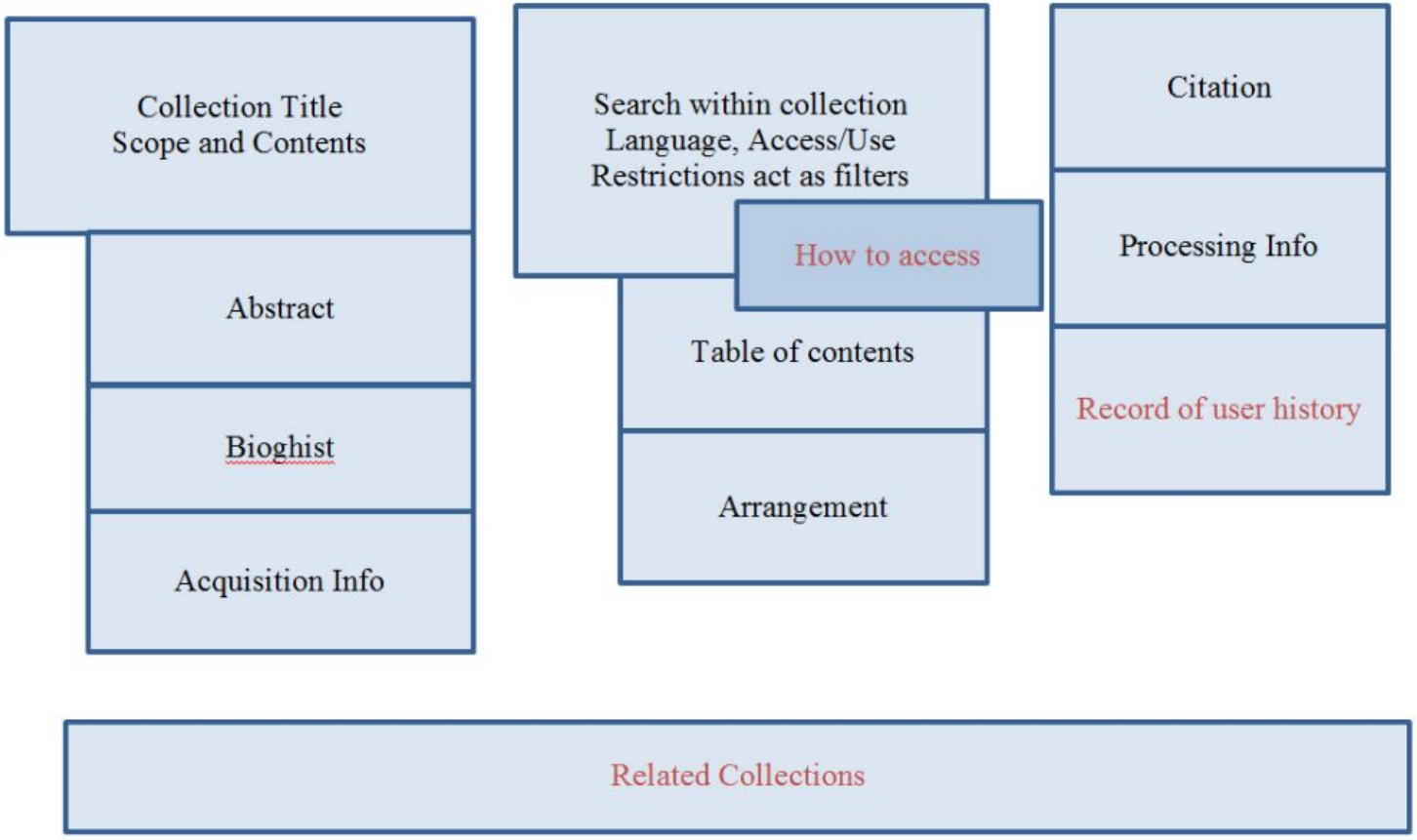
Unmoderated task-based usability testing

Guided Site Interviews

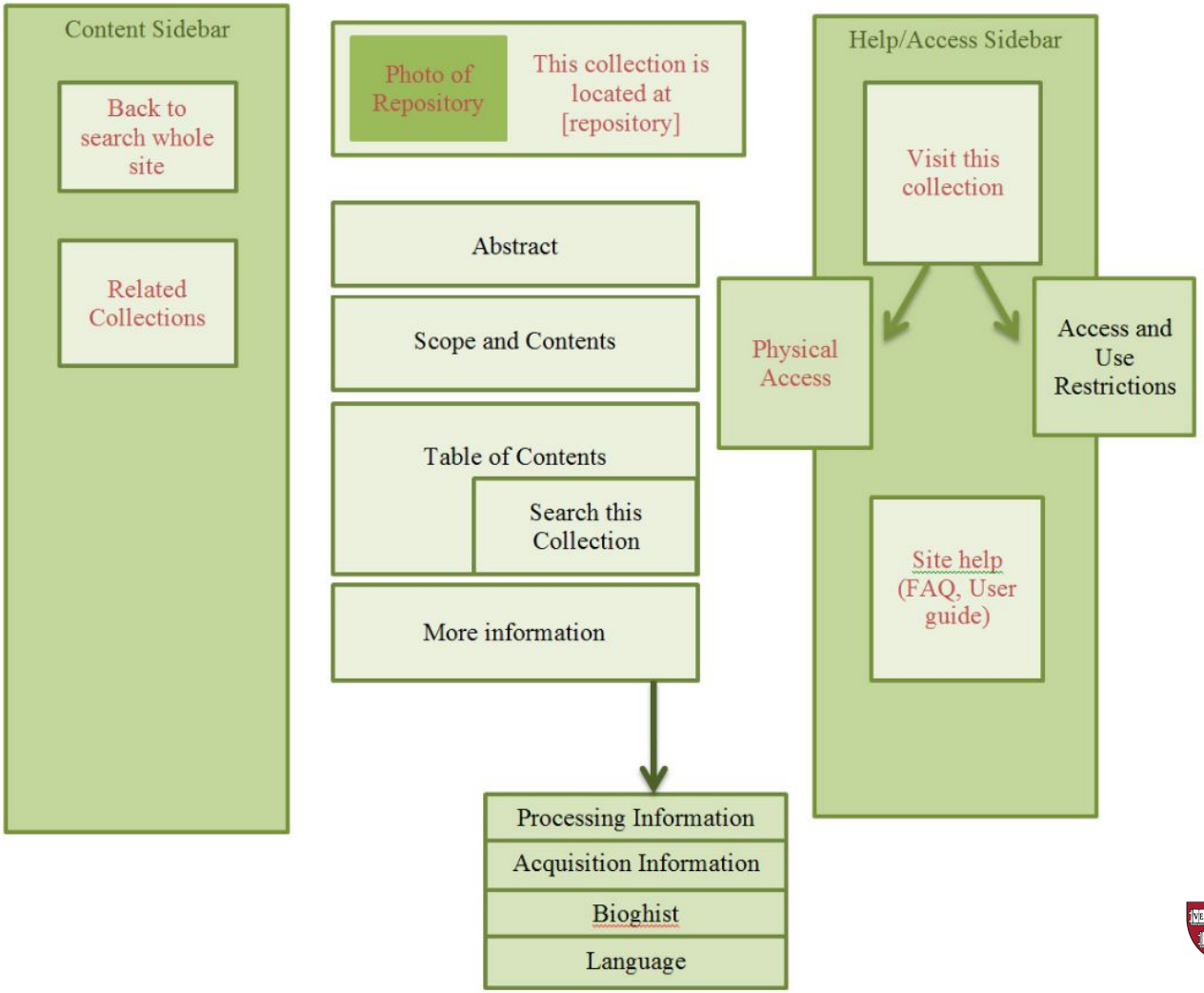
Card Sort

Modified Card Sort/Vocabulary-generating exercise

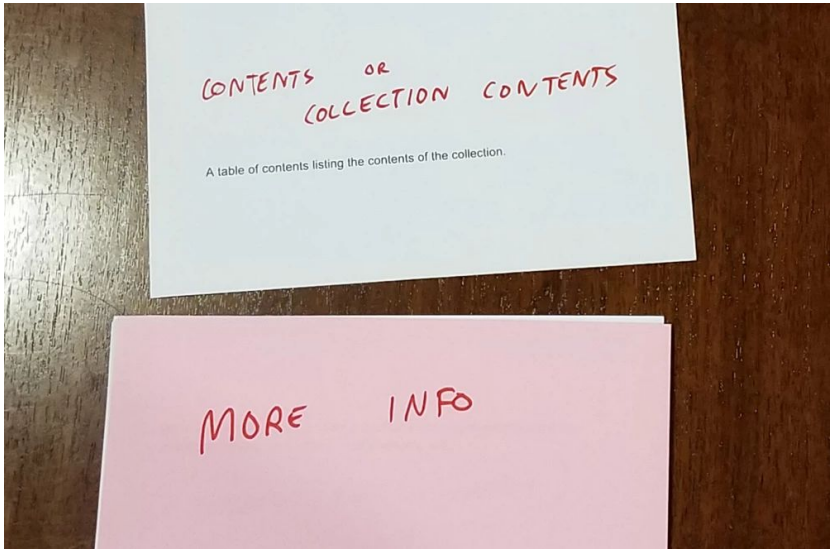
HARVARD LIBRARY



HARVARD LIBRARY



HARVARD LIBRARY



From Recommendations:

For future development, it may be useful to think about ways that users can interact with the data in ways that offer them more personal and directly useful ways to sort and resort, filter, expand, and further explore.

HARVARD LIBRARY

ArcLight

illuminating archives

Mark A. Matienzo / @anarchivist / 28 July 2017

Collaboration & Interoperability Architect, Stanford University Libraries

For more information, visit http://bit.ly/arclightproject

Description and objectives

- Project initiated by Stanford University in 2014 to address a long-standing interest in discovery/delivery of information in archives
- Support discovery of physical and digital objects
- Compatibility with and intended for integration with other systems, e.g. ArchivesSpace and Hydra-based repositories
- Development, enhancement, and maintenance by the Blacklight/Hydra communities
- **Maintain a community focus throughout the project**

The ArcLight design process

- Process led by 2 UX designers in Stanford Libraries' Digital Library Systems & Services group (Gary Geisler and Jennifer Vine)
- Followed a model for **user-centered design** developed and refined over time (see DLF 2014 presentation)
- **Community-oriented, collaborative design process** was an intentional choice and existed from the beginning
  - Informed by Stanford's participation in open source projects
  - Opportunity for other institutions to identify needs and participate in work
  - Build interest and identify potential commitments for software development

Design process contributors

Individuals

Gary Geisler, Stanford University Libraries
Wendy Hagenmaier, Georgia Tech
Nabeela Jaffer, University of Michigan
Mark Matienzo, Stanford University Libraries
Sarah Newhouse, Chemical Heritage Foundation
Kayla Ondracek, University of Michigan
Chris Powell, University of Michigan
John Rees, National Library of Medicine
Mike Shallcross, University of Michigan
Jennifer Vine, Stanford University Libraries
Laura Wilsey, Stanford University Libraries

Institutions

Bentley Historical Library (University of Michigan)
Chemical Heritage Foundation
Columbia University
Georgia Tech
Getty Research Institute
Indiana University
National Library of Medicine
New York University
Pennsylvania State University
Rockefeller Archives Center
Stanford University
United States Holocaust Memorial Museum
Yale University

Design process components

Discovery

Phase 1 → Phase 2

Environmental scan
Stakeholder goals
Interview planning

Archivist interviews
End-user interviews
Interview analysis

Information Architecture

Phase 1 → Phase 2

User needs
User personas
Requirements prioritization

Conceptual models
Conceptual sitemaps
Wireframes

Visual Design

Phase 1

Visual design mockups
Visual design style guide

Tracing user needs

- Idea/need: provide delivery of digital objects in context of description
- Environmental scan: investigate existing implementations and encoding practices
- Interview questions
  - Archivists
    - What proportion of [your repository's collections] are digital archives or have some digital content?
    - Do you have requirements for linking finding aids to digital content? Do you currently do this? How?
    - How would like to see links to digital objects instantiated? In other words, if your finding aid links to a digital image, how would that digital image be displayed? What about other media types?
  - Researchers
    - How important to you is access to the digital content in an archival collection?
    - Do you choose an archival collection based on whether it includes digital content?

Tracing needs: interviews

Sample quotes from interview analysis on user needs

ArcLight, to the extent that it's feasible, we want to give the user access to the digital objects within ArcLight… Pretty straight forward for something like images, even for video…Even an embeddable viewer…at some point. **(Archivist)**

I think the distinctions [where content is stored] are important for us … for knowing where things are, especially if there is a reprocessing project or we need to verify something. That's what we use our current collections management database for… But probably I feel researchers don't care where things live as long as they can have access to them **(Archivist)**

ArcLight probably can't have an embeddable viewer, accomplish delivery of every file format. The great thing about finding aids, is we have this link. For the most part, for crawls and [digital repository], [you] probably get a better view. **(Archivist)**

[Access to digital content is] super important if I can't come to the archive. Even if I have the money to go to an archive, I'm only going to look at the stuff that isn't digitized. **(Researcher)**

The other issue, I do think when things are digitized it's easy for them to… get the sense that they're not in a continuum next to other things or in folders or together in a way. Sometimes, it's very crucial how things are or left or either reorganized or whatever it is. The things that are nearby. **(Researcher)**

http://bit.ly/arclight-design-documents

Tracing needs: personas

ArcLight
Administrator Persona

Marcia Garza
Systems Archivist
Archival Collections department at an independent research institution

After working as a junior web developer during grad school, Marcia is enjoying her first year working as an archivist. The work is not without its frustrations, however, and she's excited enough about the potential of ArcLight to improve access to the archive's collections that she volunteered to be their first ArcLight administrator. She's more tech-savvy than most of the other archivists here, but she's not a programmer and she's a bit nervous about whether she'll be in over her head.

"I'm more than willing to spend whatever time it takes to configure ArcLight if it will tame the confusion and chaos that currently limits access to our fantastic collection of archival materials."

Frequency of ArcLight use: Daily

Technical proficiency: High. Marcia suffers from repetitive stress disorder and prefers to use only a keyboard.

	Motivation	Scenario	Specific Goals
1	Easily customize ArcLight to be institution-specific	When she logs into the ArcLight administration page, Marcia is pleased to find it even more straightforward than the Drupal-based websites she's worked on. She quickly uses a few forms to change the basic appearance of the site to reflect the institute's name, logo, and colors. It doesn't take much more effort to update the site's footer, to specify the recipients of the feedback form, and to customize the template of a virtual reading room consent form to reflect the Institute's policies. Despite the efforts of the Institute's public service archivists, Marcia has noticed that many researchers don't understand how to effectively use the Institute's archives or how to request items from the collections. To help address these issues, Marcia creates support pages that outline her Institute's reading room policies, explain how to make digitization requests, and describe the Institute's local conventions and descriptive practices.	• Customize labels and appearance options • Add support pages with institution policies, etc. • Customize an integrated feedback form • Customize a template for a virtual reading room consent form
2	Selectively activate and configure desired features	Marcia had concerns about how well ArcLight would accommodate some of the idiosyncratic ways the Institute describes and makes archival materials discoverable. As she works through the Configuration section of the administrative interface, she finds that many features can be turned on or off, and some features can be customized. She can select to expose only some of the available social media features, for example, associate specific content viewers with different types of content, and make some detailed choices about how search results are displayed. She can also enable and customize some options related to full-text search of digital objects. She sees that some defaults can be overridden at the collection level so that an archivist can determine whether the settings are appropriate for a given collection.	• Selectively enable features of interest, such as specific social media features • Configure details of complex features, such as content viewers, full-text search, and search results display
3	Configure viewers by content type	When she first set up ArcLight, Marcia accepted the default viewers for all digital content types. The support developer for her department has recently created and installed a new email viewer that is closely integrated with the Institute's authentication service. Marcia updates the configuration to use the new viewer for all email content. She gives the heads-up to an archivist that a web-based text-mining plugin created by an ArcLight community member has been installed. The archivist will want to update one of her recent collections to make the writer's Word documents available to this plugin. As a rule, though, archivists don't need to revisit collection descriptions to update viewers: once Marcia sets a default for a content type, it applies to all objects of that type automatically.	• Configure viewers for rendering digital material within (or one-click-away-from) ArcLight • Archivists don't need to revisit collections to make sure the viewers are up to date

ArcLight
Advanced Researcher Persona

Anna Chandler, Ph.D.
Associate Professor
Department of American Culture at a large public university

Dr. Chandler's research and teaching interests are centered in the Arab and Muslim American Studies Program. She uses physical and digital collections at the archives to introduce her students to primary source materials, and for her own publications and conference presentations. She is experienced with various public access catalogs, finding aid interfaces, and online collections. She frequently takes digital photographs of archival materials with her cell phone, and uses Zotero to manage citations.

"As I'm engaged in both teaching and research, I have to make the most of my time when I'm at the archives. I browse finding aids and access online content when I'm at home or my office while preparing for class or conducting research."

Frequency of ArcLight use: Weekly

Technical proficiency: Moderate

	Motivation	Scenario	Specific Goals
1	Find primary source materials relevant to pedagogical and research interests	Dr. Chandler is beginning work on an article about the contemporary Muslim-American experience in Detroit. She begins by doing a keyword search across the entirety of the archives' holdings, then facets the results by geographic location and date range. The search results include collections, specific collection components, and digital objects. She selects a new collection, the Midwest Ecumenical Forum record, opens its finding aid, and starts to browse its intellectual hierarchy before conducting several keyword searches within the collection. She bookmarks a few items of possible interest. She returns to the initial results, and selects a relevant folder title from the Nabeel Hassan papers. When the component opens, she sees that it is part of the collection's "Islamic Center" subseries (which in turn belongs to a "Topical Files" series) and that there are no restrictions on access or use of the materials.	• Employ multiple search/discovery strategies across and within collections • Search full text where available • Receive search results that include collections, components, and digital content • Facet and refine search results • Browse/navigate the hierarchy of an archives
2	Request items or digitization in advance of visits to the archives	Dr. Chandler has identified a potentially interesting folder in the Nabeel Hassan papers. She clicks the "Request This" button next to the folder title so that the material will be ready for her next visit to the reading room, and then continues to review the finding aid for relevant content. As she peruses the scope and content note for a series of "Articles, Manuscripts, and Poems," she realizes that a volume of Hassan's diary will be essential to her research; she therefore goes to the component in ArcLight and clicks the "Digitization Request" button. By the time she returns to her initial search results, she is alerted to two notifications in her account: the box she requested will be ready for her in two days (and will be held at the reading room for a week) and her digitization request has been queued for processing.	• See clear and unambiguous information regarding applicable use or access restrictions • Submit item requests and digitization orders • Receive notifications about requests • Save collections, components, digital objects to a bookmark list
3	Gain direct access to digital content (as permitted by access/use restrictions)	Dr. Chandler facets her search results to show only digital content and discovers a broad range of relevant information in multiple content types. She notices that items associated with a particular email account in the Khalil Alawan papers are restricted to reading room-only access due to copyright issues and makes a note to view them on her next trip to the archives. Moving on, she finds a number of oral histories from Muslim community leaders in the Jawad Mazrui papers that she can stream remotely in ArcLight via an embedded player. She then turns her attention to some born-digital office files in the Charles Hassan papers that are only accessible to members of the Midwest University community. After clicking on the item and authenticating using her registered username and password, she gains access to a virtual reading room. She navigates through the directory structure and is able to render common file formats in-browser via an embedded player.	• Access digital content: directly from finding aid/discovery interface; from various sources (Hydra, Archive-It, ePadd, etc.); in-browser as much as possible • Easily find information regarding access/use restrictions (including copyright) • Access the virtual reading room for materials with access/use restrictions

ArcLight
Archivist Persona

Coretta Rey
Arrangement and Description Archivist
Special Collections department in a large public university

With 12 years of experience, Coretta is an expert at producing archival metadata. She's passionate about making archival collections discoverable and accessible, but frustrated with workflows that require her to produce description in multiple systems, export to several others, and wait days or months in between. As her repository's backlog of unprocessed hybrid and born-digital collections grows, Coretta is excited about exploring new ways to provide access to digital materials in a user-friendly discovery environment.

"I want to get archival information to the user easily, quickly, and completely—giving researchers streamlined access to all sorts of digital material within ArcLight, as much as possible."

Frequency of ArcLight use: Daily

Technical proficiency: Moderate to high

	Motivation	Scenario	Specific Goals
1	Easily publish and update archival descriptions and digital collections	Coretta just finished processing a large hybrid collection—the papers of a well-known writer. She created description in ArchivesSpace, and used Forensic Toolkit Imager, BitCurator, Archive-It, and Archivematica to process and package the born-digital materials and ingest them into her Fedora/Hydra repository. In ArcLight, she selects her data sources, and selects the indexing options she wants. She knows researchers will want to search the content of the writer's drafts and correspondence, not just her description, so she selects full-text indexing. When the indexing is complete, she previews what the collection will look like to the end-user. She catches an error in the description, and corrects it in ArchivesSpace; pushing the small update to ArcLight is quick. When she's happy with the preview, she clicks "Publish" to make the collection discoverable in ArcLight.	• Create and update archival description in any tool; index and publish it in ArcLight • Workflow optimized for integration with ArchivesSpace • Integrate digital material from different sources • Make indexing choices based on content • Preview the collection in context
2	Link digital content to the description at any level	The writer's collection includes over 2,000 digital photos. Coretta creates description for the bulk of these images at the series level, rather than the item level, knowing that in ArcLight the images will be presented as a grid of thumbnails for the user to browse or sort by embedded metadata. For a few important images the writer has annotated (ones used for covers of the writer's books), Coretta creates (a separate series with) item-level description to make them more easily discoverable. Most of the writer's manuscripts are in Word files, which display as static documents by default. Coretta's Support Developer recently installed a web-based text mining plugin created by an ArcLight community member; Coretta overrides the default viewer so the Word files open in the plugin.	• Integrate digital material into archival description at item, folder, series, and/or collection level • Override default viewer configuration for a collection • Expose embedded content metadata to end users
3	Configure discovery and access restrictions	The writer has stipulated that a series of drafts can only be accessed in the reading room, but when Coretta demonstrates ArcLight's virtual reading room functionality, he consents to virtual access. Coretta configures access so that researchers must request permission to view the drafts online, and once approved, must agree to terms of use before each access. She specifies that all emails in the collection are embargoed for 5 years; when this period expires, the emails will be automatically released in ArcLight.	• Support access restrictions at the item, folder, series, and/or collection level • Enforce embargo until the specified date has passed • Allow access to specific users/groups
4	Enhance the archival description via linked data	The archives received this writer's collection largely due to its relationship to other collections in the archives. At the donor's request, Coretta has taken the extra step to create a rich agent record, including bio notes from multiple sources, and links to other writers and collections. She would love to see network graphs and other visualizations of these rich connections in the discovery environment.	• Support EAC or other linked-data connections to other collections and contextual information • Display graph or other visualizations of these connections

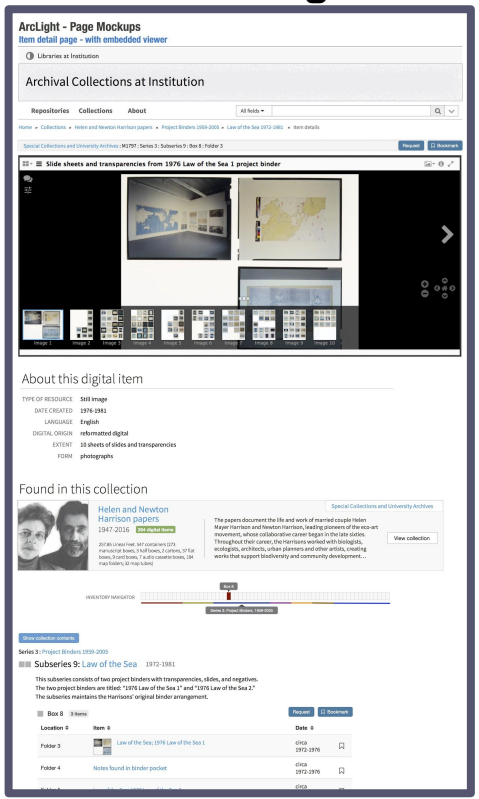
http://bit.ly/arclight-design-documents

Tracing needs: requirements

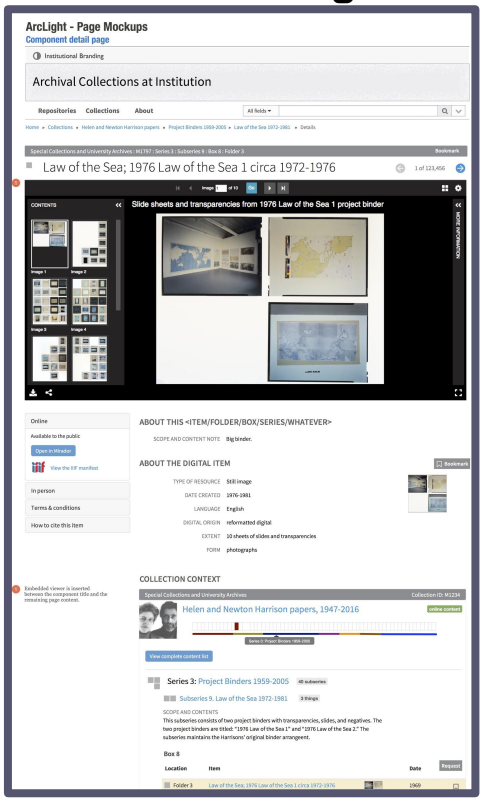
1.45	1. Core Discovery	UI features	Display/link digital material at various levels: item, folder, series, and/or collection	1 - Must have
1.61	1. Core Discovery	UI features	When components of archival collections and digital objects are presented, display a core set of descriptive and administrative metadata (including collection, series, sub-series, item, etc.) to maintain context and provenance of materials	1 - Must have
6.8	6. Digital objects	Integrated delivery	Display of AV in context of description	1 - Must have
6.9	6. Digital objects	Integrated delivery	Display of images in context of description	1 - Must have
6.13	6. Digital objects	Integrated delivery	Support for user access to digital content	1 - Must have
6.15	6. Digital objects	UI features	Clearly communicate the level (collection, series, sub-series, item, etc.) of description for digital objects	1 - Must have
6.16	6. Digital objects	UI features	Display metadata about the digital object that comes from the archival component (not necessarily the digital object), and does so in a way that allows for a predefined portion or all the metadata to display	1 - Must have
6.19	6. Digital objects	UI features	Preview digital content with thumbnails	1 - Must have
1.13	1. Core Discovery	Discovery	Bring together elements of the archival collection that might be in different silos (i.e., a Hydra repository, an Archive-It web archives collection, email in ePADD, etc.).	2 - Should have

Tracing needs: wireframes

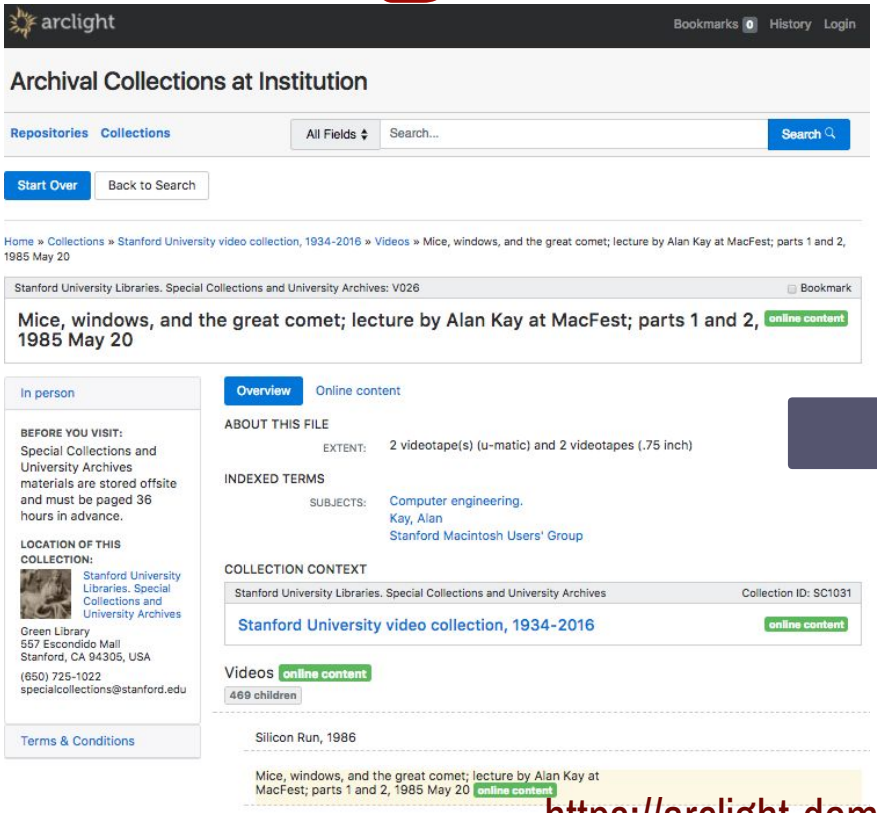
Initial design

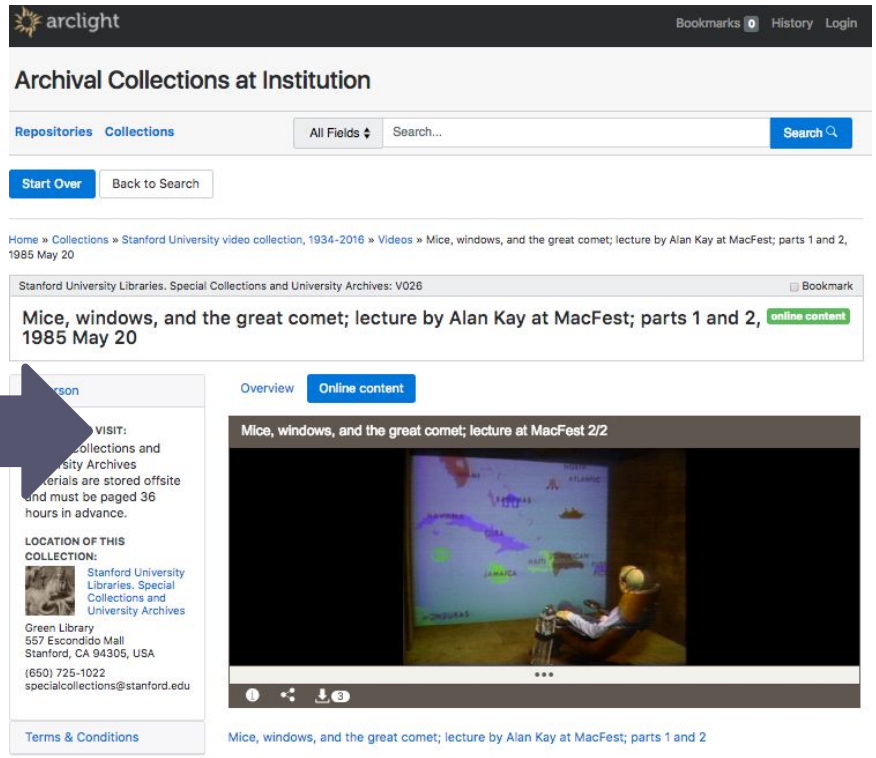


Revised design



Tracing needs: implementation





https://arclight-demo.projectblacklight.org/

Observations

- Valuable to get early and frequent input on designs
- Overhead of collaboration sometimes slows things down; requires additional coordination
- Some parts of design process require deep UX or design expertise
- Fear of design by committee mitigated by demonstrating leadership while making adequate space for critical feedback
- Being heavily community-focused allowed for greater engagement and built more excitement

Thank You!

Mark A. Matienzo / @anarchivist / 28 July 2017

Collaboration & Interoperability Architect, Stanford University Libraries

For more information, visit http://bit.ly/arclightproject

Group discussion ground rules

- Only one person should speak at any given time
- Move up/Move up (switch your speaking and listening habits)
- No one knows everything; together we know a lot
- We can’t be articulate all the time
- Education, not argumentation
- Be mindful of time
- Discussion held under Chatham House Rule: https://www.chathamhouse.org/about/chatham-house-rule

Discussion Questions

State of archival discovery

- How would you describe archival discovery at your current institution?
- What works well with your current archival discovery implementation? What helps your users to discover archival materials?
- What are the weaknesses with your archival discovery implementation or what do you want to improve? What challenges do users face in discovering archival materials?

User centered design and user assessment

- How do you see using user assessment/collaborative work as a way to address that?
- What are the challenges to doing that in your institution?
- What are the weaknesses of using a user-centered approach?
- How do you think about your users? Who are and aren't your users? Who are you designing for?
- Are there tools/techniques that would make it easier to lower the barrier to using user-centered design processes?

Collaborative work

- How has your department collaborated with other departments or institutions in the past?
- What are the barriers and opportunities you see to working collaboratively across institutions?

Google folder for notes documents: http://bit.ly/s403-discussion