This project was made possible in part by the Institute of Museum and Library Services, through grant LG-246349-OLS-20. The IMLS is the primary source of federal support for the nation's libraries and museums.

https://confluence.ucop.edu/display/NAFAN
Creating, Maintaining, and Sharing Archival Descriptions: The Experiences and Values of US Archivists

Why Are Users Searching Archives Online? Results of a User Survey from 12 State and Regional Archival Finding Aid Aggregators (NAFAN)

The Broadest Possible Range of Users: Learning About Users of Archival Aggregations

Merrilee Proffitt (she/her/hers)
Senior Manager OCLC Research
@merrileeiam / proffitm@oclc.org

Lesley A. Langa, Ph.D. (she/her/hers)
Associate Research Scientist
@lesleylanga / langal@oclc.org

Chela Scott Weber (she/her/hers)
Senior Program Officer OCLC Research
@partly_cloudy / weberc@oclc.org
# Project Team

<table>
<thead>
<tr>
<th>Organization</th>
<th>Role</th>
<th>Key contacts</th>
</tr>
</thead>
</table>
| University of California CDL | California Digital Library: Project coordination and administration. | • Günter Waibel (Principal Investigator; Executive Director and Senior Vice Provost)  
• Rachael Hu (User Experience Design Manager)  
• Adrian Turner (Senior Product Manager)  
• Esther Verreau (CDL Senior Software Developer) |
| Chain Bridge Group           | Chain Bridge Group: Business and market research, sustainability and governance model recommendations. | • Raym Crow (Managing Partner)                                                                                                                                 |
| OCLC                         | OCLC: Research to gain a current understanding of researcher and contributor needs; assess finding aid data quality. | • Chela Scott Weber (Senior Program Officer)  
• Itza Carbajal (Associate Researcher)  
• Lesley A. Langa, Ph.D. (Associate Research Scientist)  
• Lynn Silipigni Connaway, Ph.D. (Director of Library Trends and User Research)  
• Merrilee Proffitt (Senior Manager)  
• Brittany Brannon (Research Support Specialist)  
• Brooke Doyle (Project Coordinator)  
• Janet Mason (Executive Assistant to the Vice President) |
| SHIFT Collective              | Shift Collective: Project management support and community engagement activities. | • Bergis Jules  
• Erin Glasco  
• Lynette Johnson  
• Jon Voss                                                                                                                                                  |
| University of Virginia Library | University of Virginia Library: Technical assessment and prototyping; support development of specs/requirements for a minimum-viable instantiation of the system post-2022. | • Daniel Pitti (Director of the Social Networks and Archival Context Cooperative)  
• Keith Bossert (Programmer/Developer)                                                                                                                     |
Research Advisory Group

- Bill Stingone, New York Public Library
- Derek Mosley, Auburn Avenue Research Library on African American Culture and History, Atlanta-Fulton Public Library
- Hillel Arnold, Rockefeller Archives Center
- Holly Mengel, University of Pennsylvania, Kislak Center (PACSCL)
- Lydia Tang, Lyrasis
- Molly Bruce Patterson, Rhode Island College (RIAMCO)
- Rachel Walton, Rollins College
- Ricky Punzalan, University of Michigan, School of Information
Focus Group Interviews with Archivists

- 10 Interviews / 52 Archivists total
- Multiple archive types, positions ranging from entry level to leadership

Analysis focused on:
- Enabling and constraining factors related to creating archival description
- Enabling and constraining factors related to participating in aggregation
- Perceived value of aggregation
Research Data Analysis

Two major sources of user data:

• Pop-Up survey
  • Hosted on 12 Regional Aggregator sites
  • 3300 complete, usable responses
  • Gathered demographic and research needs info
• Semi-structured individual interviews
  • 25 interviews, 45-60 min
  • 5 user types: Academic, Advanced Academic, Professional, Personal Interest, Genealogist

Analysis focused on:

• Describing and classifying users, based on needs, motivators, behaviors
• Benefits and challenges of using aggregators
CREATING, MAINTAINING, AND SHARING ARCHIVAL DESCRIPTIONS: THE EXPERIENCES AND VALUES OF US ARCHIVISTS
Focus Group Interview Summary

• 10 focus groups
• 52 participants total
  – 24 participating in aggregation
  – 28 not participating in aggregation
  – Representing 27 states and District of Columbia
• Divided into five group types
  – Leadership / Administrative
  – Public Services & Tech Services, small institution
  – Public Services & Tech Services, large institution
  – Technical Services & Systems Responsibilities
  – Very Small Institution / Lone Arrangers
Participants by Archive Type

<table>
<thead>
<tr>
<th>Archive Type</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic archive</td>
<td>22</td>
</tr>
<tr>
<td>Museum library / archive</td>
<td>11</td>
</tr>
<tr>
<td>Government archive</td>
<td>6</td>
</tr>
<tr>
<td>Public library archive</td>
<td>4</td>
</tr>
<tr>
<td>Community archive</td>
<td>4</td>
</tr>
<tr>
<td>Historical society archive</td>
<td>3</td>
</tr>
<tr>
<td>Independent research library</td>
<td>2</td>
</tr>
<tr>
<td>International organization</td>
<td>1</td>
</tr>
<tr>
<td>Performing arts archive</td>
<td>1</td>
</tr>
<tr>
<td>Regional archive</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>52</strong></td>
</tr>
</tbody>
</table>

Carnegie Classifications Breakdown

- 2 worked at archives in Baccalaureate Colleges (1 Arts & Sciences Focus, 1 Diverse Fields)
- 5 worked in Master’s Colleges & Universities (2 M1, 3 M2)
- 15 worked in Doctoral Universities (10 R1, 4 R2, 1 D/PU)
Participants by US Region

- **WEST**: 11 total
- **MIDWEST**: 8 total
- **NORTHEAST**: 13 total
- **SOUTH**: 20 total

The map shows the distribution of participants across different US regions, with specific numbers indicated for each region.
Aggregation Contributor Research Questions

• What are the **enabling and constraining factors** that influence whether organizations describe the archival collections in their care?

• What are the **enabling or constraining factors** that influence whether organizations contribute to an aggregation of archival description?

• What **value** does participation in an archival aggregation service **bring** to organizations?
Final Protocol

• Please tell us about your role in working with archival collections at your institution.
• Please tell us briefly about how archival description gets produced and published in your institution.
• This question focuses on producing and publishing archival description where you work. What are the things that make producing and publishing archival description in your home institution easy or difficult? What are the things that support this work and make it easier, either that you have now or would like to have?
Final Protocol (continued)

• Think about your current process / the hypothetical scenario where your home institution would contribute to an archival aggregation platform. What about the process is/would be easy? What about the process is/would be difficult?

• What makes/would make participation in a finding aid aggregation valuable for your institution? What makes/would make your participation in a finding aid aggregation valuable for your researchers? Do you think there are other things an aggregator could offer that would be valuable to your institution or researchers?

• If you had a magic wand and could create your ideal national-scale finding aid aggregation system, how would it work and what would it do?
Four Themes

• Things that constrain creating archival description
• Things that enable creating archival description
• Incentives to participate in aggregation (current or perceived)
• Disincentives to participating in aggregation (current or perceived)
<table>
<thead>
<tr>
<th>Theme</th>
<th>Sub-Theme</th>
<th>Sub-sub Theme</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Things that constrain creating archival description <strong>Circumstances</strong>, resources, or realities that prevent the archivist or institution from creating archival description. Use these codes only for the creation of archival description, not to capture what is said about sharing or contributing description to an aggregator/aggregation.</td>
<td>a. Complex workflows/workarounds People are using multiple systems and steps to get work done. Include where people are dealing with MARC records, PDF, etc.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>b. Creating EAD is a challenge Include where people say they are not creating EAD because it is too hard and when they are doing EAD but it is hard. Includes references to the complexity of EAD.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c. Revised description Descriptions need to be revised and revisited for numerous reasons on an ongoing basis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>i. legacy description <strong>Revising</strong> existing description, includes both older full description and minimally processed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ii. reparative description <strong>Anti-racist description</strong>, addressing harmful language, “decolonizing”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>iii. Other types of description that</td>
<td></td>
</tr>
</tbody>
</table>
Constraints

• Complex workflows / workarounds (n=29)
• Dealing with legacy description (n=19)
• Lack of control over IT / software (n=16)

Enablers

• Collection Management Systems (n=21)
• Worksheets / templates (n=14)
• Documentation / manuals (n=8)

n=number of focus group interview participants
Constraints

“...our library IT is baffled by our needs.”
-Special Collections Archivist, Academic archive, West

Enablers

“ArchivesSpace makes it easy for us...It's been a sea change for us in terms of time savings.”
-Supervisory Archivists, Academic archive, West
Incentives to participate

• Increased visibility / awareness of collections (n=25)
• Uncovering connections between collections (n=22)
• Ease of contributing records \textit{in an automated way} (n=21)

Disincentives to participate

• Data requirements (n=19)
• Difficult to edit / update records (n=14)
• Difficult to submit records (n=8)

\textit{n=number of focus group interview participants}
Incentives

“...we're so small and such a niche [topical archive] that we're not a destination necessarily for a lot of people.”

-Archivist, Independent research library, West

Disincentives

"We can't delete things. It is a process for us to delete, we have to put in a request and make sure that it gets taken down...We're doing an anti-racist descriptive practices deep-dive into all of our finding aids, we're having to do this whole process of uploading and editing over and over again."

-Special Collections archivist, Academic archive, West
Description constrainers

• Complex workflows / workarounds (n=29)
• Dealing with legacy description (n=19)
• Lack of control over IT / software (n=16)

Description enablers

• Collection Management Systems (n=21)
• Worksheets / templates (n=14)
• Documentation / manuals (n=8)
WHY ARE USERS SEARCHING ARCHIVES ONLINE?
RESULTS OF A USER SURVEY FROM 12 STATE AND REGIONAL ARCHIVAL FINDING AID AGGREGATORS (NAFAN)
Who are the current users of finding aid aggregations? Do they align with the persona types and needs identified in recent archival persona work?

What are the benefits and challenges users face when searching for descriptions of archival materials within finding aid aggregations?
Data that describe need

- **Demographics:** where are researchers located? How old are they? What is their profession?

- **Purpose for researching:** Why are they using the aggregation site?

- **Behaviors:** How frequently do they visit the site? Do they use other sites for their search? If yes, which ones?
Survey Methods

Purpose
- Capture data on users of online aggregators
- Identify potential interviewees for qualitative data collection

Design
- Hosted by 12 aggregators
- Convenience sample
  - Opt-in
- Opened March 18 - Closed on May 31
- ~3,300 responses
Survey Questions

- Demographics
- How the user found the aggregator
- Purpose for research
- Type of material wanted for research
- Preference for digital vs. physical materials
- Impact of COVID on research
Pop-up Users, by profession

- Retiree: 20.8%
- Archivist/Librarian: 13.5%
- Graduate/Postgraduate Student: 9.6%
- Faculty/Academic Researcher: 9.2%
- Genealogist: 8.4%
- Professional: 6.5%
- Undergraduate Student: 5.9%
- Journalist/Writer: 4.9%
- Arts/Artist/Filmmaker: 4.9%
- Lifelong Learner: 3.7%
- Museum Professional: 3.1%
- K-12 Education: 2.1%
- Historian: 2.1%
- Independent Researcher: 1.5%
- Unknown: 1.0%

N=3,078
Pop-up Users, by project purpose

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal Interest</td>
<td>32.7%</td>
</tr>
<tr>
<td>Family History Research</td>
<td>24.4%</td>
</tr>
<tr>
<td>Professional Project</td>
<td>22.8%</td>
</tr>
<tr>
<td>Long term Project</td>
<td>20.0%</td>
</tr>
<tr>
<td>Local History Research</td>
<td>19.3%</td>
</tr>
<tr>
<td>Class Assignment</td>
<td>7.2%</td>
</tr>
<tr>
<td>Creative or Artistic Project</td>
<td>6.5%</td>
</tr>
<tr>
<td>Thesis or Dissertation</td>
<td>5.0%</td>
</tr>
<tr>
<td>Material to Use for Teaching</td>
<td>4.7%</td>
</tr>
<tr>
<td>Short-term Project</td>
<td>4.1%</td>
</tr>
</tbody>
</table>

N=3,172

*Percentages do not sum to 100%
Top 5 professions by project purpose

*Percentages do not sum to 100%
How Pop-up Users found aggregator websites

- I found it through a browser search (Google, Yahoo, Bing, etc.): 44.0%
- I followed a link on a website or social media: 20.6%
- I've used this site before: 20.3%
- Other: 7.0%
- Someone told me I'm about this site: 5.8%
- Not sure how I got here: 2.1%
How the top 5 professions found aggregator websites

- Archivist/Librarian: 14.0% Social media, 7.1% Other
- Faculty/Academic Researcher: 14.0% Social media, 8.7% Other
- Genealogist: 21.1% Social media, 4.7% Other
- Graduate/Postgraduate student: 17.5% Social media, 8.7% Other
- Retiree: 21.5% Social media, 6.6% Other

Less than 4% for all users for "Not Sure"
How frequently pop-up users visit aggregator websites

- This is my first time, 55.4%
- Daily, 2.8%
- Weekly, 9.0%
- Monthly, 12.1%
- Less than monthly, 16.3%
- Other, 4.5%

N=3,172
Top 5 professions visit frequency to aggregator websites

- Archivist/Librarian: First time, 18.8%; Weekly, 22.0%; Monthly, 21.3%; Less than monthly, 17.3%
- Faculty/Academic Researcher: First time, 40.5%; Weekly, 8.0%; Monthly, 17.1%; Less than monthly, 26.4%
- Genealogist: First time, 61.8%; Weekly, 5.3%; Monthly, 9.7%; Less than monthly, 17.6%
- Graduate/Postgraduate student: First time, 46.5%; Weekly, 15.6%; Monthly, 12.7%; Less than monthly, 19.6%
- Retiree: First time, 74.3%; Monthly, 6.2%; Weekly, 2.5%; Less than monthly, 12.4%

Less than 2% for all users except Archivists for daily use
Less than 1% for all users for "Other"
### Pop-up Users, by material preference

<table>
<thead>
<tr>
<th>Material</th>
<th>Preference (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any material relevant to my topic</td>
<td>55.8%</td>
</tr>
<tr>
<td>Photos</td>
<td>39.3%</td>
</tr>
<tr>
<td>Personal Family Papers</td>
<td>37.2%</td>
</tr>
<tr>
<td>Government Records</td>
<td>31.1%</td>
</tr>
<tr>
<td>Magazines</td>
<td>26.0%</td>
</tr>
<tr>
<td>Oral Histories</td>
<td>22.6%</td>
</tr>
<tr>
<td>Books</td>
<td>21.1%</td>
</tr>
<tr>
<td>Maps</td>
<td>20.2%</td>
</tr>
<tr>
<td>Corporate Records</td>
<td>14.8%</td>
</tr>
<tr>
<td>Videos</td>
<td>11.6%</td>
</tr>
<tr>
<td>Art</td>
<td>10.7%</td>
</tr>
<tr>
<td>Audio</td>
<td>10.1%</td>
</tr>
<tr>
<td>Other</td>
<td>10.1%</td>
</tr>
<tr>
<td>Finding Aids</td>
<td>1.3%</td>
</tr>
<tr>
<td>Architectural papers</td>
<td>0.6%</td>
</tr>
</tbody>
</table>

**N=3,172**

*Percentages do not sum to 100%*
TwoStep Cluster Method

Cluster Quality

Silhouette measure of cohesion and separation

- Poor
- Fair
- Good

Cluster Breakdown:

- Archivists/Librarians & Other Professionals: 23.4%
- Faculty & Academic Researchers: 18.4%
- Undergraduate & Postgraduate Students: 10.8%
- Family History Researchers: 23.8%
- Personal Interest Researchers: 23.5%
NAFAN User Clusters

- **Archivist/Librarian & Other Professionals (N=559)** – public and technical service archivists/librarians, journalists, writers, tradespeople

- **Faculty and Other Doctoral Degrees (N=328)** – faculty, post-doc, independent researcher, museum professionals

- **Family History Researchers (N=722)** – local history, professional genealogist, hobby genealogist

- **Personal Interest Researchers (N=711)** – general search with a range of profession and education

- **Undergraduate and Postgraduate Students (N=712)** – Indicate their profession is undergraduate/postgraduate student
THE BROADEST POSSIBLE RANGE OF USERS: LEARNING ABOUT USERS OF ARCHIVAL AGGREGATIONS
• Extant studies focus on the needs of academics, especially historians and other humanists (Rhee, 2015)

• Literature investigating broader cross-section of archival users indicates that different researcher types have significantly different needs. (Duff, 2005)

• Archival software projects have developed personas reflecting a more diverse spectrum of researchers (ArcLight, 2017; Project Electron, 2019; ArchivesSpace, 2016)
Sampled from pop-survey participants

25 interviews, 5 per persona type
- Archivist/Librarian,
- Faculty and Doctoral Degrees,
- Family History Researcher,
- Personal Interest Researcher,
- Undergraduates/Postgraduates
Semi-Structured, Virtual, 1-1 interviews, 45-60 minutes

1 interviewer, 1 notetaker from research team

Interviews collected July - August 2021

Codebook developed through inductive/deductive approach March - June 2022

Using QDA NVivo with a 3-coder team with coding running June - July 2022
Interview Research Questions

Who are the current users of aggregated archival description?

Do current user types align with the persona types and needs identified in recent archival persona work (i.e., what characteristics are present vs not)?

Why are the current users trying to discover and access archival collections via aggregation of archival description?

How are current users discovering and accessing aggregations of archival description?

What are the benefits and challenges users face when searching archival description in aggregation?
Interview Supplemental Research Questions

What future directions would users want archival aggregators to take?

What other tools are users using to discover or access archival materials and description?
Codebook Themes

**Discovery** - Refers to the ways users search for and identify archival resources (primary child codes: Challenges, Enablers, Benefits)

**Access** - Refers to the ways users locate, retrieve, and use archival resources (primary child codes: Challenges, Enablers, Benefits, Types of Access)

**Aggregators** – Comments specific to archival aggregator use and knowledge
Codebook Themes

**Needs** – references to what may satisfy a user's information need

**Motivators** – references to what motivates users to seek archival resources

**Decision Making** – user describes direct aspects impacting the process of making choices regarding archival research

**Research Duration** – timeframe of the user’s research that guides or influences a user’s pursuit of archival resources
Codebook Themes

Sources – mentions of specific information sources used in their research

Pandemic – instances and ways in which the pandemic impacted archival research

Magic Wand – any suggested changes, new additions, or ideal scenarios by users regarding archival aggregation systems
1. Please describe a typical research project and the research process you go through over the course of the project.

- **Needs** – references to what may satisfy a user's information need
- **Motivators** – references to what motivates users to seek archival resources
- **Research Duration** – timeframe of the user’s research that guides or influences a user’s pursuit of archival resources
2. We identified you via a survey you took at [Aggregator in Month]. What were you looking for? Are there specific reasons or situations for which you use [Aggregator]?

- **Aggregators** – Comments specific to archival aggregator use and knowledge
3. Please describe the ways you look for archival materials. What other resources and/or tools do you use to find archival materials?

- **Discovery** – Refers to the ways users search for and identify archival resources (primary child codes: Challenges, Enablers, Benefits)
- **Sources** – mentions of specific information sources used in their research
- **Needs** – references to what may satisfy a user's information need
- **Motivators** – references to what motivates users to seek archival resources
4. Once you locate the description of archival material that interests you, what do you typically do next?

- **Access** - Refers to the ways users locate, retrieve, and use archival resources (primary child codes: Challenges, Enablers, Benefits, Types of Access)

- **Decision Making** – user describes direct aspects impacting the process of making choices regarding archival research

- **Pandemic** – instances and ways in which the pandemic impacted archival research
5. Think about a recent attempt to search for archival material for your research. What did you find particularly delightful about the experience? What did you find particularly frustrating?

- **Pandemic** – instances and ways in which the pandemic impacted archival research
- **Discovery** - Refers to the ways users search for and identify archival resources (primary child codes: Challenges, Enablers, Benefits)
- **Access** - Refers to the ways users locate, retrieve, and use archival resources (primary child codes: Challenges, Enablers, Benefits, Types of Access)
6. If you had a magic wand, what would be your ideal way of finding archival material for your research?

- Magic Wand – any suggested changes, new additions, or ideal scenarios by users regarding archival aggregation systems
Interview Research Challenges

1. Persona types heavy overlap or lack of clarity in distinctions between previously identified roles.

2. Users struggle to recognize or remember specific archival aggregator sites.

3. COVID-19 pandemic restrictions introduce extraordinary circumstance behaviors and responses (from both user & institution).

4. Users don't cleanly delineate between archival and non-archival research or formats.
1. Users utilize a variety of non-aggregator sources to locate archival materials.

2. Decision making factors play a key role in how and if users pursue archival resources (in person or online).

3. Inequitable access to collections stems from structural and professional issues.

4. Many users are aware of resource challenges and limitations of archival institutions.
Works Cited


THANK YOU!  ANY QUESTIONS?