

UC Office of the President

CDL and Partner Organizations - Project Publications

Title

Building a National Finding Aid Network: Final Report

Permalink

<https://escholarship.org/uc/item/6xr6q38m>

Authors

Turner, Adrian

Schiff, Lisa

Mitchell, Catherine

[et al.](#)

Publication Date

2024-08-08

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>



University of California

CDL

California Digital Library

Building a National Finding Aid Network

Final Report

August 2024

A decorative pattern at the bottom of the page consisting of overlapping geometric shapes in shades of orange, blue, yellow, and pink.

August 2024

Report prepared by:

Adrian Turner, Lisa Schiff, Catherine Mitchell, and Günter Waibel



Permalink: escholarship.org/uc/item/6xr6q38m

This work is licensed under a Creative Commons Attribution 4.0 International License.

<https://creativecommons.org/licenses/by/4.0/>



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. To learn more, visit www.ims.gov.



Table of Contents

Note on Terminology	4
Executive summary	4
1. Introduction	7
2. Context and background	8
2.1. The Problem: barriers accessing archival materials	8
2.2. Goals and objectives of the project	10
3. Research and community engagement findings	13
3.1 Discovering and accessing archival collections	15
3.2 Responsible and community-determined access	16
3.3 Resourcing, participation, and sustainability	16
3.4 Finding aid formats and description conventions	16
3.5 Finding aid publishing infrastructure and workflows	17
4. Envisioning the network: sustainability and governance	18
4.1. Community-based membership model	18
4.2. Assessing the feasibility of a community-based membership model	20
4.3. Community-based governance model	21
5. Envisioning the network: feature and functionality requirements	23
6. Next phases	27
6.2 Building the network together	29
6.3 Conclusion	30
7. Acknowledgements	31
Appendix A: Requirements for the network	32
Overview	32
Functions	32
Data models	41
Appendix B: Assessment of existing systems	47

Note on Terminology

The terms "archives" and "finding aids," which appear throughout this document, are specific terms of art in the domain and were thus utilized within the context of this project (embedded, even, in the project name itself). For the purposes of this report, the term "archives" represents a broad range of cultural heritage and memory organizations and encompasses any entity that collects and stewards the organically-created records of individuals, families, or other organizations. The term "finding aid" refers to collection guides, inventories, registers, and other summaries of descriptions of materials represented in archival collections, intended to support researchers with context for and access to those collections. While this report reflects the terminology used at the time of the project, any future developments stemming from this project should consider more expansive terms.

Executive summary

This report is a key deliverable from the 2020-2023 "[Building a National Finding Aid Network](#)" (NAFAN) research project, and outlines a plan for building and launching a comprehensive national finding aid network that is community-driven, -sustained, and -governed.

The NAFAN project explores a fundamental challenge faced by researchers of all types: the significant barriers to locating relevant archival materials across the vast, distributed, and unevenly resourced field of cultural heritage institutions. Archival records are primarily discoverable via finding aids, which provide descriptions and summaries of the contents of individual archival collections. Yet, many archives don't or can't publish their finding aids online – and state and regional aggregations of finding aids, which are designed to increase the discoverability of these materials, often support a siloed and limited range of archives, and are at-risk as their infrastructure ages and budgets dwindle. Consequently, much of the stewarded archival content in the United States remains relatively invisible, and the voices documented therein are poorly represented in the historical record. Complicating this situation further, some archives may seek to prioritize service to their local communities rather than increasing the visibility of their collections for a broader audience.

The project addresses this discoverability challenge by laying the groundwork for a national finding aid network available to all contributors and researchers. The California Digital Library (CDL), Chain Bridge Group, OCLC, Shift Collective, and the University of Virginia Library (UVA) collaborated to design such a network via concurrent work streams: community building and broad stakeholder engagement; a research agenda to gain a current understanding of the needs of both researchers and archives in relation to finding aid aggregation; technical assessments of potential systems to support network functions; and business and governance modeling, along with market research, to address long-term sustainability from the outset.

Research and community engagement findings from the project highlight the value of a national-level network that attends to the discoverability needs of both researchers and archives. The findings reveal, in particular, the need for sustainable, centralized infrastructure to increase the capacity and visibility of archives and their collections: Researchers of all types are interested in using archival materials, but are

hampered by siloed discovery systems and must expend significant time, effort, and personal resources to locate and access relevant materials; Many archives, particularly community-based ones, lack infrastructure and resources to describe and share information about their collections; Archives have finding aids in a range of formats, prepared using a range of descriptive conventions, and when sharing finding aids with aggregators, they are consequently burdened with additive workflows that tax already minimal resources. The research additionally underscores the importance of centering local priorities, from ensuring responsible, community-determined access to archival collections to including community-based archives in the planning, design, and implementation of such large-scale projects from the outset.

The findings have fundamentally informed this proposal for designing, building, and launching a national-level network. The value of the network is ultimately tied to how relevant it can be to all US-based archives: broad, community-based participation, representation, resourcing, and governance must be at the core of this effort to align the network with the values of the larger community of individuals engaged in cultural heritage work. The project proposes a collective, community-based "ability to pay" membership model to sustain a national network once built and launched. This model realizes the efficiency and reduced costs of a shared infrastructure and is predicated on coordinated community participation. Preliminary market research findings confirm the viability of such a community-based support model. In tandem with a community-based membership model, the project recommends that the governance of the network also be embedded within the community – facilitating a long-term vision that centers equity, inclusion, and the needs of all institutions participating in and represented by the network.

The core features and functions of the network would support the broadest range of contributing organizations and minimize barriers to participation in order to:

- Scale discovery and access to a national level that is as comprehensive as possible for the broadest possible range of researchers.
- Enhance the value of an individual archive's holdings by situating their repository profiles and finding aids within a national network of related finding aids from academic, community-based, government, historical societies, and public archives – complementing each institution's holdings and amplifying their research value.
- Support archives with determining the level and degree to which they would like to participate and share information about their holdings, based on their capacity, current workflows, and preferences for the extent of information they choose to provide.
- Facilitate seamless sharing of finding aids in existing formats not limited to EAD, with connectors to platforms such as ArchivesSpace, Access to Memory, PastPerfect, and more.
- Enable technical capacity for archives while saving cost, by offering a long-term hosting and publication platform for institutions lacking local infrastructure.
- Cultivate broader capacity for archives by providing support, training, tools, and funding resources for archives through a peer-based network.

This report recommends a continued phased, incremental, and community-driven approach that moves this effort from a research project to a robust program over a 4-5 year horizon. As a community-driven initiative, NAFAN will offer community members many opportunities to engage in design, planning, and implementation activities across the next three phases of the project, following the completed Phase 1 research work:

- **Phase 2:** Establishing a coalition; expanding project leadership; identifying potential long-term administrative/service hosts, and funding partners (~1-2 years).
- **Phase 3:** Developing and launching a "Minimum Viable Product" (MVP) version of the network; garnering community support (~2 years).
- **Phase 4:** Expanding MVP functionality to support a more fully-featured network and governance structure (~1 years).

The archives community has an imperative to provide richer and more comprehensive access to collections held by a diverse array of US-based archives – documenting a far broader and representative set of communities, voices, and topics than are currently discoverable. And we can do this in a more sustainable way than we have managed in the past, by abandoning siloed, duplicative infrastructure and leveraging a large-scale, national finding aid network that is community-driven, -sustained, and -governed. This is a timely and overdue opportunity for us to coalesce to meet shared challenges. NAFAN could have a transformative and lasting impact on archives and the researchers they serve by reducing barriers to discovery, expanding the historical record, and offering a new path to sustainable aggregation for archives nationwide.

1. Introduction

This report is a key deliverable from the 2020-2023 "[Building a National Finding Aid Network](#)" (NAFAN) research project, rooted in the goal of providing inclusive, comprehensive, and persistent access to finding aids by laying the foundation for a national network available to all contributing archives and researchers. This document is primarily intended for archivists, curators, and other memory workers (including staff, volunteers, managers, and directors) who are stewarding and providing access to holdings within US-based archives ranging from historical societies, museum archives, and community-based archives to academic research library archives, government archives, and public library archives.

The report outlines a plan for building and launching a network in subsequent project phases, supported by a broad coalition of partner archives, researchers, and other stakeholders – particularly ensuring that community-based archives and other non-academic research library archives are at the table from the start.

This report includes the following sections:

- Context, goals, and objectives of the project
- Findings from the project's research and community engagement activities
- A proposed community-based sustainability and governance model for the network, with resourcing estimates
- High-level feature and functionality requirements for the network
- Recommendations for next steps to realize the vision for the network
- Appendices with detailed descriptions of the requirements and supporting data models, and an assessment of existing technical platforms

The goal of the report is to introduce and socialize a proposed model for realizing a national finding aid network, that would provide significant value both to archives and to researchers. The intent behind sharing this work is to cultivate interest, engagement, and advocacy towards realizing this vision.

2. Context and background

Researchers of all types face significant barriers to locating relevant archival materials across the vast, distributed, and unevenly resourced field of cultural heritage institutions: Many archives don't or can't publish their finding aids online; State and regional aggregations of finding aids are siloed and support a limited range of archives; A good number of these aggregations are also at risk as their infrastructure ages and budgets dwindle. Consequently, much of the stewarded archival content in the United States remains relatively invisible, and the voices documented therein are poorly represented in the historical record. Complicating this situation further, some archives may seek to prioritize service to their local communities rather than increasing the visibility of their collections for a broader audience. In response to these challenges, the NAFAN research project explored how more inclusive, comprehensive, and long-term access to archival materials can be achieved by advancing a national finding aid network that is community-driven, -sustained, and -governed – meeting archives "where they are at" based on their capacity, current workflows, and preferred access models. In order to inform the vision for this network, the California Digital Library initiated a research agenda to gain a current understanding of the needs of both researchers and archives with regard to finding aid aggregation.

2.1. The Problem: barriers accessing archival materials

The NAFAN research project was initiated to address a fundamental challenge faced by researchers of all types: the significant barriers to locating relevant archival materials across the vast, distributed, and unevenly resourced field of cultural heritage institutions.

The limited visibility of archival collections

Archival collections are primarily discoverable via finding aids which provide descriptions and summaries of the contents of individual archival collections. These critical scholarly resources enable students, teachers, genealogists, historians, and a broad array of researchers to discover, locate, identify, and understand archival resources, especially those in analog formats. Unfortunately, finding aids themselves are often difficult to find: many institutions do not publish their guides online. While many archives want to make their finding aids broadly visible on the open web, they may lack resources and infrastructure to display them on their own platforms – with varying success in making them visible to a broad community or to search engines.

The relative invisibility of finding aids renders many archival collections themselves invisible to all audiences, and presents a compounding inequity for those scholars without funding to travel to archives to explore the materials in depth.

Regional aggregation as a partial solution

Over the past two decades, aggregators have formed at the state and regional level, with significant early investments by institutions and funders, and now number 12 across the country.¹ These aggregators collect and provide access to finding aids on behalf of multiple institutions, and have historically made strong inroads towards addressing some of the discoverability challenges described above. Aggregators provide additional core services (e.g., free tools, training, and support) to help institutions create and share finding aids in standard formats. Many aggregators also serve as the de facto or primary publication venue for those institutions that lack the capacity to maintain their own publication system. In all cases, aggregators provide additional visibility for the finding aids they collect.

Despite the work of these aggregators, however, researchers still lack broad and persistent access to archival materials via finding aids: not all US regions have a finding aid aggregator, and the aggregators that do exist maintain siloed platforms that display finding aids from institutions only within a limited service area. As a result, researchers remain unable to comprehensively explore US-based archival resources across more meaningful dimensions, such as provenance or subject matter.

Compounding this problem, thousands of geographically eligible archives are still not represented within extant aggregations. Many organizations have finding aids and collection descriptions in formats unsupported by most aggregators, such as spreadsheets, MS Word, HTML, and PDF. As a result, countless archives – particularly small, lesser-resourced institutions – are less visible on the open web, and the voices documented in their collections are poorly represented in the historical record.

The limits of the existing ecosystem

Apart from the significant challenge of providing broader access to distributed archival materials, many aggregators struggle even to find sufficient resources to maintain the status quo, let alone invest in improvements; with aging systems and negligible budget lines, they are fundamentally at risk, which is in turn a major risk for the archival community and researchers around the world. The dissolution of aggregators would spell the loss of many online finding aids where the aggregator has served as the institution's primary publication platform – and could cost the broader archives community both resources and users. Any attempts at institution-based solutions would, by definition, result in greater duplication of effort and an even more scattered landscape of finding aids for researchers.

Complementary platforms such as [Social Networks and Archival Context \(SNAC\)](#) and [ArchiveGrid](#) also play a role in the national landscape, leveraging and indexing finding aid and other data gathered from aggregators and individual institutions – and serving as an additional access point for researchers to identify relevant archival holdings. However, SNAC and ArchiveGrid themselves are dependent on the

¹ As of August 2024, this includes: Archival Resources in Wisconsin, Archival Resources of the Virginias (ARVAS), Archives West, Arizona Archives Online (AAO), Black Metropolis Research Consortium (BMRC), Chicago Collections Consortium, Connecticut's Archives Online (CAO), Empire Archival Discovery Cooperative (Empire ADC), Online Archive of California (OAC), Philadelphia Area Archives, Rhode Island Archives and Manuscripts Online (RIAMCO), and Texas Archival Resources Online (TARO).

persistent access to finding aids made available by aggregators (and individual institutions). All these challenges necessitate that we rethink aggregation at a national scale.

Reconsidering the access model

This research project started with the assumption that there is a universal desire for broad access to archival materials; this perspective, however, was challenged and ultimately refined through discussions with community-based archives, some of whom made clear that access by their own community members is their primary, if not sole, focus. Broad visibility may also not align with the need of archives to provide responsible access to collections, where the safety and well-being of individuals and communities represented in the collections must be ensured. These findings refine our understanding of the space and have informed the design of NAFAN's proposed next phases, from governance to feature recommendations, as described in the remainder of this report.

2.2. Goals and objectives of the project

The above challenges, though significant, have presented an opportunity to rethink aggregation at scale – based on a clear understanding of the needs of researchers and archives in relation to finding aid aggregation, supported by research findings rather than current assumptions or the status quo.

Validating the vision of an inclusive national finding aid network

In 2018-2019, the California Digital Library (CDL) organized a planning initiative, "[Toward a National Archival Finding Aid Network](#)," supported by the Institute of Museum and Library Services (IMLS) under the provisions of the Library Services and Technology Act, administered in California by the State Librarian. The CDL convened a symposium of US finding aid aggregators and expert advisors with deep knowledge of organizational development, community engagement, and sustainability to explore a proposed new aggregation model. Preparatory research and outcomes from the symposium clearly confirmed that current aggregations are insufficient in their coverage of archival materials and are at-risk in their current siloed contexts.²

Symposium participants called for a substantially more inclusive framework to meet the needs of a more diverse set of end users, enable participation by a wider range of archives, and support a greater variety of collection description levels and formats, while simultaneously transitioning away from outmoded technologies and directly addressing foundational issues of sustainability. They conceived of this as a national finding aid network that is community-driven, -sustained, and -governed, and formulated an action plan outlining a concrete path toward this set of goals. The vision for this network comprised a suite of shared infrastructure and services supporting:

² Allison-Bunnell, J. (2019). *Finding Aid Aggregation at a Crossroads*, <https://escholarship.org/uc/item/5sp13112>. See also 2018-2019 Background Resources on the NAFAN research project wiki at <https://ucop.edu.atlassian.net/wiki/spaces/NAFAN/pages/47743740/Project+Reports+and+Resources>

- Meaningful, inclusive, and low-barrier pathways to participation by any archive in the US, accepting formats beyond [Encoded Archival Description \(EAD\)](#) and [MARC21](#) and providing tools and support to easily create and publish finding aids;
- An extensive and continually updated registry of archives, regardless of whether or not they have finding aids, to increase the awareness among researchers of organizational profiles and holdings; and
- A primary or supplemental publication platform (as needed by the contributing archive) with comprehensive aggregation of and persistent access to finding aids, as well as integrations with related additional contextual information (e.g., biographical and historical information about persons, families, and organizations that created or are documented in archival collections) and digital versions of materials described in the finding aid.

Building the future and designing the network

The NAFAN research project, funded through a grant from the Institute of Museum and Library Services ([grant LG-246349-OG-20](#)), was initiated from this action plan, and was coordinated by the CDL, in collaboration with [Chain Bridge Group](#), [OCLC](#), [Shift Collective](#), and the [University of Virginia Library](#) (UVA), and in close partnership with [statewide/regional finding aid aggregators](#). [ArchivesSpace \(LYRASIS\)](#) also served as a technical consulting partner.

The project focused on four concurrent activity streams:

- **Community building and broad engagement** with aggregators, individual archives, and other stakeholders to support engagement with the project and future development and implementation of the network.
- **Research** to understand the needs of both archival researchers and archives regarding finding aid aggregation, and for an evaluation of the quality of existing finding aid data within aggregations.
- **Technical assessments** of potential systems to support network functions, and the formulation of system requirements for a minimum viable product instantiation of the network.
- **Business and governance modeling, along with market research**, to identify viable sustainability and governance strategies for the network once built and launched.

Each collaborator had a specific focus within the project: Chain Bridge Group led preliminary business research activities, and formulated sustainability and governance model recommendations; OCLC led research activities; Shift Collective supported project management and community engagement activities; and UVA led technical assessment and prototyping activities. A core project team (including representatives from the CDL, Chain Bridge Group, OCLC, Shift Collective, and UVA) administered the project activities, with support from three consultative and advisory groups:

- An Aggregator Partners Working Group of program managers from regional/statewide finding aid aggregations,
- A Research Working and Advisory Group of advisors knowledgeable about research methods for surfacing the needs and practices of end-users and archivists, and

- A Technical Working and Advisory Group of technical domain specialists to provide input and direction on technical assessment and prototype development.

With the support of Shift Collective, the project team identified a need to articulate [core values and principles](#) to guide project activities and to undergird the formulation of recommendations outlined in this report. Shift Collective also surfaced the importance of proactive and specific engagement with community-based archives, many of whom are not represented in regional or national professional archival organizations, to explore their needs in relation to finding aid aggregation – and to that end convened specific discovery and feedback sessions with representatives from such organizations. Shift Collective also helped the project transition to hosting openly-available information sessions as opposed to presenting updates at conferences with definitionally limited participants, a strategy reinforced by the suspension of all in-person activities due to the COVID pandemic. This approach also included the cultivation of a mailing list and public Slack channel to share key project deliverables. While the pandemic necessitated a transition to the virtual from the in-person, the team realized that this new approach was also essential for achieving broader and more open opportunities for community engagement.

3. Research and community engagement findings

The project's research and community engagement findings highlight the value of a national-level network that attends to the discoverability needs of both researchers and archives. The findings reveal, in particular, the need for sustainable, centralized infrastructure to increase the capacity and visibility of the archives community and its collections: Researchers of all types are interested in using archival materials, but are hampered by siloed discovery systems and must expend significant time, effort, and personal resources to locate and access relevant materials; Many archives, particularly community-based archives, lack infrastructure and resources to describe and share information about their collections; Archives have finding aids in a range of formats, prepared using a range of descriptive conventions, and when sharing finding aids with aggregators, they are often burdened with additive workflows that tax already minimal resources. The research additionally underscores the importance of centering local priorities, from ensuring responsible, community-determined access to archival collections to including community-based archives in the planning, design, and implementation of such large-scale projects from the outset.

A foundational core value of the NAFAN research project was to eschew assumptions and have the realities of the archival research ecosystem be the driver of all proposed future work.³ To that end, the project began with a foundational investigation of the needs of the wide variety of stakeholders in this space. This work, led by OCLC and Shift Collective, was undertaken to surface the diverse perspectives of both researchers and archives, and to ground the requirements and design of NAFAN in those real goals and aspirations.

OCLC was responsible for conducting and synthesizing research to understand the needs of both archival researchers and archives regarding finding aid aggregation. OCLC also evaluated the quality of existing finding aid data as represented within regional aggregations, to explore what could be done with existing data within a national-level aggregation context and to also lay the groundwork for iterative data remediation and expanded network features over time. Data gathering and analysis followed a mixed-methods approach, including: a survey placed on the sites of 12 regional archival aggregators, yielding a total of 3,352 usable responses from researchers describing their search behavior, information needs, and demographic characteristics⁴; semi-structured individual interviews with 25 researchers to explore their information needs and information-seeking behavior⁵; focus group interviews with 53

³ "We must base our long term plans for the network on our research findings rather than current assumptions or the status quo"; see the NAFAN research project wiki:

<https://ucop.edu.atlassian.net/wiki/spaces/NAFAN/pages/47744025/Our+Goals+Values+and+Principles>

⁴ Langa, L. A, Weber, C., & Connaway, L. (2023). *Pop-up Survey: Findings from the Building a National Archival Finding Aid Network Project*, <https://escholarship.org/uc/item/5qp9z6tb>

⁵ Weber, C., Carbajal, I. A, Langa, L. A, Connaway, L., Doyle, B., Brannon, B., & Proffitt, M. (2023). *User Interviews: Findings from the Building a National Archival Finding Aid Network Project*, <https://escholarship.org/uc/item/7hn3m08f>

archivists to examine their needs for describing collections and contributing finding aids to aggregators⁶; and a quantitative analysis of over 145,000 EAD encoded finding aids, collectively representing 741 repositories.⁷ These findings have been synthesized in a 2023 report, *Summary of Research: Findings from the Building a National Archival Finding Aid Network Project*.⁸

In tandem with the research activities conducted by OCLC, Shift Collective staged a two-day workshop in December 2021, inviting participants whose work is centered on community-based cultural memory projects. The workshop aimed to explore the kinds of services, support, and advocacy that a national network could provide to community-based archives, in order to help researchers better discover and access collections stewarded by these organizations. The findings were summarized in a 2022 report, *Increasing Discovery of Archives: A Project to Provide Better Pathways to Archival Records in Cultural Heritage Collections*.⁹

A strong theme across the research and community workshop findings is the significant role a national network could play in increasing visibility and discoverability of archival materials, to the extent desired by a given contributing archive – while also helping to address the inequities that exist within current models of access to archival materials. Per OCLC's findings:

"Alignment between the needs of end users and archives illuminates how a national aggregation can provide value by addressing the urgent shared needs around the visibility and discoverability of archival collections. These include addressing researchers' difficulty in discovering archival collections on the web and accessing the archival materials they identify and archives' challenges in creating and publishing archival description and making their archival collections visible on the web. Our research also identifies a significant disparity in the resources available to both archives and archival researchers, which impacts the historical record, scholarship, and knowledge access and dissemination. Addressing the broad needs of disparate archives and archival researchers would be an important intervention in an uneven access landscape."¹⁰

These findings are echoed in Shift Collective's report, which underlines a more systemic need for equitable resourcing, advocacy, and peer-based networks to support the health of the broader archives ecosystem – especially for under-resourced community-based archives. The workshop findings also highlight a critical caveat, reshaping a core assumption articulated at the outset of this project about the benefits of broad discoverability: enabling broad and open availability of finding aids does not necessarily equate to connecting collections to specific primary audiences and communities supported

⁶ Weber, C., Proffitt, M., Langa, L. A., Connaway, L., Brannon, B., & Doyle, B. (2023). *Focus Group Interviews: Findings from the Building a National Archival Finding Aid Network Project*, <https://escholarship.org/uc/item/81s9129f>

⁷ Washburn, B., Proffitt, M., & Weber, C. (2023). *EAD Analysis: Findings from the Building a National Archival Finding Aid Network Project*, <https://escholarship.org/uc/item/90k6w2wf>

⁸ Weber, C., Connaway, L., Doyle, B., Langa, L. A., Proffitt, M., Washburn, B., & Carbajal, I. A. (2023). *Summary of Research: Findings from the Building a National Archival Finding Aid Network Project*, <https://escholarship.org/uc/item/3tv1m5xj>

⁹ Shift Collective. (2022). *Increasing discovery of archives: A project to provide better pathways to archival records in cultural heritage collections*, <https://escholarship.org/uc/item/0kd7d5m0>

¹⁰ *Summary of Research: Findings from the Building a National Archival Finding Aid Network Project*, p.iv

by archives. The increased visibility can place added pressures on already-strained public service operations. Broad and open access must also be tightly aligned with processes and protocols enabling institutions to provide responsible access (which can sometimes mean restricted access) to the collections they steward, to protect the safety and well-being of individuals and communities represented in the collections.

A number of specific themes emerged across all of the research and community engagement findings, grouped and summarized in the sections below.

3.1 Discovering and accessing archival collections

- The central challenge for researchers is the insufficient discoverability of archival collections on the web: researchers are hampered by having to use siloed discovery systems and expending significant time and effort locating relevant materials, without certainty that they have identified all of the appropriate sites to explore. Researchers and archivists underscored the need for an easy-to-use system that provides comprehensive discovery of archival collections across many institutions.
- Researchers use a variety of technical tools and platforms to conduct and scope searches for relevant archival materials. Search engines are a key discovery tool.
- Researchers face challenges with understanding and parsing large results sets across search-based systems. They desire a way to narrow results and easily assess what is and is not relevant to their research. They are also uncertain about the comprehensiveness of their searches, not understanding the coverage of their search tools, and are concerned about missing out on newly available collections.
- A range of audiences – and not solely academic researchers and genealogists – utilize finding aid aggregations, and are seeking archival materials for a range of academic, personal, and professional uses. This includes many first-time and less frequent users of archival aggregation, who may need support in understanding and using the interface – as well archivists, librarians, and other information professionals who frequently use archives within public services contexts.
- Researchers' and archivists' expectations for discovery system features and functionality need to be balanced with the extent and quality of existing finding aid data (in both structured and unstructured formats).
- Researchers and archivists want to surface connections between collections, including options for browsing and refining searches based on headings for people, organizations, places, and topical subjects.
- Researchers continue to seek seamless access to online versions of materials described within finding aids.
- Researchers need pathways to access physical archival materials once they have discovered them – and noted the ways aggregators, catalogs, websites, and search engines currently support or frustrate that process.
- Researchers described how resource limitations impact their ability to conduct in-person research, or even to license or request copies of materials, as many avocational and academic researchers rely on personal funds to subsidize their research.

3.2 Responsible and community-determined access

- Enabling broad and open availability of finding aids needs to be fundamentally aligned with processes and protocols to support institutions with providing responsible access to the collections they steward – ensuring "the safety, security, privacy, and dignity of communities represented in those collections".¹¹
- Some community-based archives are most concerned with addressing the need, not yet fully met, to ensure that local community members can access their collection descriptions, irrespective of whether academic researchers or a national audience are able to.
- Increased visibility of collections puts a greater demand on community-based archives' public service functions, particularly licensing and usage requests.
- Respecting cultural heritage and cultural identity protocols is vital. A conundrum for some community-based archives is that making their collections more accessible can open the door to cultural appropriation and misuse.
- Archives need and value control of data in an aggregation environment, in order to share finding aid data and/or repository information on their own terms. This also includes sharing of their data with third-parties.
- Archives are seeking to improve the quality of finding aids by remediating harmful language and using self-descriptions from communities; updates to finding aids need to be propagated from their versions of record through to aggregations.

3.3 Resourcing, participation, and sustainability

- Archives desire robust system functionality and support services from a national-level network.
- Archives of all sizes want a network that supports and includes small and under-resourced institutions, including at low or no cost for the latter.
- Community-based archives are acutely underfunded and lack basic resources. Many are consumed with sustaining their organizations and operations at a bare-bones level and require support to develop local capacity to process, describe, safely house, and steward their collections for the long-term.
- The network must provide multiple types of value for key stakeholders: archivists, administrators of archival programs, and researchers. Contributing organizations need information, analytics, and impact data for understanding and communicating the value of their efforts.

3.4 Finding aid formats and description conventions

- Archives maintain and publish a broad range of finding aid formats not limited to EAD, including PDF files, text documents, and spreadsheets.

¹¹ *Increasing discovery of archives: A project to provide better pathways to archival records in cultural heritage collections*, p. 20

- Community-based taxonomies and local descriptive conventions play important roles that could be obscured by mandating a particular descriptive standard within the context of an aggregation.
- Archivists and researchers expressed unmet desires and expectations related to discovery that could be at least partially met with additional structured data.
- Existing EAD-based finding aids hosted by aggregators contain markup that highly correlates to *Describing Archives'* (DACS) recommendations for a "Single-Level Optimum" record, affording a baseline level of structured data that could be leveraged to better meet researcher needs.¹²
- Despite conforming to a widely used structure, existing EAD-based finding aids contain inconsistent information such as the contact details for an archive, which could benefit from standardization and remediation.

3.5 Finding aid publishing infrastructure and workflows

- Many archives still lack resources and core infrastructure for describing and providing online access to their collections.
- Archives wish to be able to share their finding aids using low-friction workflows with a low barrier to participation.
- In cases where institutions have implemented archival management systems, those platforms have eased processes associated with creating and publishing finding aids.
- Participating in an aggregation is typically burdened with a range of additive and cumbersome required workflows and in particular, those specific to generating EAD outputs optimized for the aggregator.

These findings simultaneously reinforced and challenged the assumptions held going into the NAFAN research project, broadening and deepening the initial understanding of the potential value of aggregation in the larger archives ecosystem.

At a more structural project level, critically surfaced was the importance for any future efforts to include from the outset community-based and other non-academic research library archives in the planning, design, and implementation of activities – while also ensuring that all partners are equitably compensated for their participation.¹³ Collectively, these findings fundamentally inform the proposed community-based sustainability and governance models ([Section 4](#)), feature and functionality requirements ([Section 5](#)), and next phases ([Section 6](#)).

¹² Society of American Archivists (2023). *Describing Archives: A Content Standard*, <https://saa-ts-dacs.github.io/>

¹³ *Increasing discovery of archives: A project to provide better pathways to archival records in cultural heritage collections*, pp. 6-14.

4. Envisioning the network: sustainability and governance

The findings have fundamentally informed this proposal for designing, building, and launching a national-level network. The value of the network is ultimately tied to how relevant it can be to all US-based archives: broad, community-based participation, representation, resourcing, and governance must be at the core of this effort to align the network with the values of the larger community of individuals engaged in cultural heritage work. The project proposes a collective, community-based "ability to pay" membership model to sustain a national network once built and launched. This model realizes the efficiency and reduced costs of a shared infrastructure and is predicated on coordinated community participation. Preliminary market research findings confirm the viability of such a community-based support model. In tandem with a community-based membership model, the project recommends that the governance of the network also be embedded within the community – facilitating a long-term vision that centers equity, inclusion, and the needs of all institutions participating in and represented by the network.

This section offers a summarized proposal for instantiating a network that provides a community-sustained and -governed approach. As documented in the findings from the precursor planning phase to the NAFAN research project, "Toward a National Archival Finding Aid Network," most state and regional finding aid networks exist in a state of chronic underfunding and under-resourcing.¹⁴ Though federal, state, and private funders provided initial grants to establish these regional aggregations, they did not commit to sustaining those aggregations over time. Thus, individual institutions and/or existing regional aggregators, all operating within a resource-limited environment, are poorly positioned to instantiate and sustain a national network on their own.

In response to these limitations, NAFAN proposes a collective, community-based "ability to pay" membership model to sustain a national network once built and launched. This model offers the efficiency and reduced costs of a shared infrastructure and is predicated on coordinated community participation. In tandem with a community-based membership model to sustain the network, the governance of the network should also be embedded within the community – facilitating a long-term vision that centers equity, inclusion, and the needs of all institutions participating in and represented by the network.

4.1. Community-based membership model

The key proposed elements of the community-based membership model are summarized below. Importantly, the successful deployment of this model is a core requirement for sustaining the network over time.

¹⁴ *Finding Aid Aggregation at a Crossroads*, p.7.

Initial build and launch of the network

To build trust and engagement within the community, the initial design, build-out, and launch of an "Minimum Viable Product" (MVP)¹⁵ version of the network should be subsidized by grants from funding partners, supporting the participation of early-adopter partner organizations across the spectrum of archives (as discussed further in [Section 6](#), "Phase 3"). The MVP will function as a pilot to demonstrate the value of the network to the broader archives community, enabling them to evaluate the system and make future financial contribution commitments. Given sufficient up-front membership pledges, the further build-out and deployment of a fully-featured network and associated governance structure would begin, informed by community feedback during the MVP pilot phase and still largely grant-funded (as discussed further in [Section 6](#), "Phase 4").

Sustaining the network

Once the fully-featured network has been built and launched, the ongoing maintenance of the service should be wholly supported through an "ability to pay" membership model – largely predicated on the broad participation by institutions with stable funding lines. This model has the following elements:

- Membership is available to individual archives, with an option for consortia or aggregators to participate via a "bundled" discounted membership fee, where the consortia or aggregator serves as the intermediary.
- Membership tiers are centered on an "ability to pay" approach, with tiers established on the relative budget capacity of the individual archive, rather than the organization's host institution, e.g., a university campus. A free option must be available for organizations that lack stable funding lines and cannot afford membership costs. As a starting point, tier levels could be based on the number of permanent FTE (as a proxy for operating budgets).
- Any funding commitments over and above what is needed to support core operations costs would be allocated to 1) advocacy and capacity-building programs for under-resourced archives and 2) a development reserve for staffing resources to support large-scale feature implementations or platform upgrades.¹⁶

Resourcing the program

The ongoing maintenance and incremental improvement of a robust, fully-featured network, once launched, will require dedicated staffing and program resources. As discussed in [Section 4.3](#), the network's dedicated program staffing should be situated within an organization that can serve as the administrative/service home and fiscal sponsor for the network.

¹⁵ A MVP constitutes a product with sufficient core features to be deployed and usable by stakeholders and users. The product can then be elaborated on and expanded in subsequent iterations, based on feedback from stakeholders and users.

¹⁶ Supplementary funds from donations, grants, and sponsorships could be leveraged for specific strategic research and development initiatives. However the ongoing maintenance of the service would not be predicated on these funding lines.

The following is a sample staffing and program profile, at a conservative baseline. This estimate was formulated by reviewing the staffing profiles for other national finding aid and archival description aggregators and by extrapolating from CDL's experience with maintaining the Online Archive of California (OAC) statewide finding aid aggregation.¹⁷

- **Permanent full-time staff**
 - **Program**
 - **Program Manager:** Provides oversight of the network, management of the membership program, and budget; coordinates strategic planning and partnerships, working/advisory groups, program assessment and reporting; and supports development/operational activities.
 - **Community Manager:** Coordinates development/operations activities and member support services (e.g., onboarding, helpdesk management, training and consultation programs); leads communication and outreach programs.
 - **Program Operations Assistant:** Provides administrative support for the network staff and program operations.
 - **Technical**
 - **Technical Lead/Developer:** Manages technical infrastructure and applications, guiding development/operations activities, and establishing and managing finding aid contribution workflows.
 - **Developer:** Support development/operations activities and implementation of finding aid contribution workflows.
- **Contract staff**
 - **User Experience / Design Consultant:** Provides user experience and design consultation services.
- **Direct costs**
 - Infrastructure hosting and storage (e.g., cloud-based services) and software/application licensing (e.g., helpdesk system).
 - Stipends for network community members who are volunteering time to participate in network governance and advisory/working groups.
 - Travel for program staff.

4.2. Assessing the feasibility of a community-based membership model

Building on the project's research findings, CDL recognized the need to further assess the feasibility of a community-based support model – and gain insight into and gauge cost sensitivities associated with an "ability to pay" membership strategy. An additional goal was to further understand the degree to which proposed features and functions of the network would provide value to individual archives.

¹⁷ This frame of reference includes staffing levels to support Archeion, ArchiveGrid, Archives Hub, Archives Portal Europe, ARCHIVESCANAADA.ca, and the Social Networks and Archival Context (SNAC) programs.

To conduct this assessment, [Delta Think](#) was contracted to coordinate preliminary market research.¹⁸ 25 directors of archives were invited to participate in market research interviews. This group spanned a range of institution types (e.g., academic research library archives, community-based archives, historical societies, government archives, independent research archives), and ensured distribution across two primary market research segment facets: 1) resourcing level, defined by the number of FTE (as a proxy for operating budgets), and 2) participation in a state/regional aggregation. The key findings from the market research are summarized below:

- **Nearly all interviewees were very enthusiastic about the key functions and features envisioned for the network.** Discovery and display of aggregated repository records and finding aids via a Research Portal garnered the most positive response. Interviewees also indicated the need for low friction workflows to contribute data. Some interviewees noted that while their institutions did not need a hosted publication platform provided by the network, they recognized the value this could offer to institutions that lack local solutions.
- **Reactions to the tiered membership model, including the presence of a free tier, were positive across all segments.** Some interviewees noted FTE sizing needed to take nuances into account (e.g., contract vs. permanent positions, degree to which FTE are specifically allocated to the department's archives programs, etc.).
- **Interviewees expressed the need to understand impact on resources, implementation costs, and details about the network's infrastructure, to make an informed determination on participating in the community-based membership model.** Archives with larger operating budgets (correlated to over 15 FTE) have the most flexibility to participate; archives with smaller operating budgets (correlated to 1-5 FTE) have the least flexibility to participate.
- **Archives currently participating in regional/state aggregations valued the prospect of a comprehensive, national-level network, while also valuing services provided by their regional aggregator.** In particular, regional training and support programs were highlighted – suggesting complementary roles for a national-level network and regional networks.

These preliminary market research findings confirm that a community-based support model is viable. Additional work is necessary to solidify membership pricing tiers, based on the latest research regarding archives-specific resourcing levels.¹⁹

¹⁸ Any long-term administrative/service host and fiscal sponsor for the network will want to conduct more extensive market research to confidently define the community-based membership tiers in order to ensure sufficient revenue to cover network operating costs.

¹⁹ While a comprehensive inventory of all US archives and their individual resourcing levels would be optimal, for NAFAN's purposes a solid starting point would be the IMLS-funded A*CENSUS II project findings, which summarized the profiles of 746 US-based archives of a range of types and sizes. For more information, see ITHAKA and Society of American Archivists (2023). *A*CENSUS II: Archives Administrators Survey*, <https://doi.org/10.18665/sr.318227>.

4.3. Community-based governance model

As previously noted, the structure and protocols of the network's governance must both ensure sustainability of the network and support open and equitable participation across participating organizations, regardless of their ability to contribute fees. This model should include individual archives of all types and sizes – as well as regional and state aggregators, development partners, and related archival networks (e.g., archival collection management systems, discovery networks). Thus, a community-based governance model should include:

- **Centering the network within an organization that can serve as the administrative/service home and fiscal sponsor for the network.** The network's dedicated program staffing should also be situated within this organization.
- **A governing board representative of the membership and broader stakeholders.** Seats should be reserved for various types of stakeholders, including: representatives elected from across the membership tiers; representative(s) from the administrative/service home and fiscal sponsor; and network program staff. Additional seats could be reserved for other stakeholders, including representative(s) from regional finding aid aggregations, organizations providing services adjacent to the network (e.g., digital collections aggregators, developers of archival collection management systems), and researchers.
- **Charter articles of association, specifying the manner in which the network's governance may evolve over time to better reflect the range of stakeholders.** These articles of association should be undergirded by a formal mission statement and statement of values and goals.

5. Envisioning the network: feature and functionality requirements

The design of the network is motivated by more than a set of technical considerations; it must encompass functional, business, and operations requirements, with adequate staffing and infrastructure resources to maintain a production-level service that is national in scope. The network must also ensure that archives can select only those components that best meet their needs, allowing them to flexibly determine the level of participation appropriate to their circumstances. Core network features include: A publicly-available Research Portal supporting the discovery and display of the most comprehensive aggregation of repository profile records and finding aids shared by US-based archives; Flexible finding aid sharing and publication workflows, supporting a range of formats beyond EAD; "Self-serve" options for contributors to manage and assess data shared via the network, supported by a Contributor Dashboard, utilities, and reporting tools; Onboarding, training, consultation, and support services for contributors, to facilitate participation in the network; Outreach, networking, and resourcing advocacy and investment programs to build capacity across US-based archives.

This section describes the essential features and functionality of the network, including program, operations, and technical elements necessary to support a robust, equitable, production-level service that can be provided to thousands of US-based archives (see [Appendix A](#) for a detailed description of each requirement and associated data models). Importantly, the system is designed to ensure that archives can select to use only those components that best meet their needs, allowing them to flexibly determine the level of participation appropriate to their circumstances.

The system fundamentally is geared towards amplifying the broad discoverability and collocation of information about archival collections, on the open web and to all audiences. However, this is a relatively coarse mechanism for connecting collections to specific primary audiences and communities supported by archives. Moreover, this amplification must be tightly aligned with processes and protocols enabling institutions to provide responsible access to the collections they steward. The proposed features and functions outline options for contributors to specify the degree and extent of information that they would like to make available about their holdings, through the repository profile records and/or finding aids they share – with clear mechanisms for specifying the terms of use, re-use, and access to the collections that they steward. This is a starting point, however; in instantiating the network, the participation and input from stakeholders is needed to define additional relevant features and functions.

Several core functions are systems-based and dependent on a specific underlying technical architecture. The project team conducted an environmental scan of existing solutions and evaluated them based on: openness of the code and the existence of community contributors; adaptability to a distributed content model; and support for the proposed requirements. (See [Appendix B](#) for a summary of the evaluation). The team's conclusion is that no readily-adaptable existing platform meets the proposed systems-based requirements for the network. The project hence recommends developing a new platform that is intentionally designed based on the proposed requirements for the network.

Discovery and display of aggregated repository profile records and finding aids

NAFAN's publicly-accessible Research Portal will serve as a discovery and display layer to the most comprehensive aggregation of repository profile records and finding aids shared by US-based archives, both ensuring broad discovery by promoting indexing of content in commonly-used search engines and providing a robust user experience for researchers. The Research Portal will be designed to support a range of audiences with varying levels of experience working with archival materials, including first-time users as well as seasoned researchers and archivists.

The Research Portal will enable full-text searching, browsing, faceted filtering, and display of all aggregated repository profile records and finding aids. In order to support a robust and uniform experience, structured collection-level descriptive information will be derived from contributed finding aids and repository profiles wherever possible. In cases where the data cannot be derived, NAFAN will request (but not require) this data from contributors.

Researchers will be able to traverse from search and browse results in the Research Portal to a landing page for each participating institution and contributed finding aid, with pathways for contacting those institutions. Finding aid landing pages will support both NAFAN and contributor hosted materials by providing a full-view option for the former and a link to the display on the contributor's website for the latter. Finding aid landing pages will also support links to associated digital versions of materials, when available.

The Research Portal will provide clear mechanisms and direct avenues for researchers and archives to report errors, web accessibility issues, removal requests (e.g., based on legal, privacy, or other considerations), and harmful descriptive information. Recognizing the growing interest in sharing and re-use of data at scale, NAFAN will also support human and programmatic reuse of material, adhering to any re-use restrictions expressed by contributors.

Registry of repository profile records

NAFAN will maintain a Registry of any archive sharing repository profile records and/or finding aids through the network. Published repository profile records will be discoverable through the Research Portal. NAFAN will provide archives with options for adding and/or editing the descriptive information in their repository profile records, as well as options for enabling publication of their repository profile records with or without associated finding aids.

Flexible finding aid sharing and publication workflows

NAFAN is intended to be as inclusive, resource efficient, and "friction-free," as possible for contributors.²⁰ Thus, NAFAN will accept finding aids in a range of formats not limited to EAD, and will establish contribution workflows for both externally-hosted and NAFAN hosted material, ensuring that all content, regardless of hosting, will be comprehensively aggregated for indexing, searching, browsing, and display within the Research Portal.

²⁰ Contributors include individual archives, as well as any regional aggregations serving as a conduit to NAFAN.

Supporting both externally and NAFAN hosted contributions requires use of a distributed content model in which NAFAN would harvest, web crawl, and/or otherwise ingest externally-hosted finding aids shared by contributors into a central repository – while also aggregating them with NAFAN-hosted finding aids that participants have opted to create, manage, and publish using NAFAN provided services. This complete set of finding aid data would then be full-text indexed and displayed within the context of the publicly-available Research Portal.

The NAFAN "distributed" model can be contrasted with the "centralized hosting" model in use by US-based regional aggregations today. While the centralized model is well-known and established, and does provide a consistent and immersive user experience and display environment for all finding aids, that consistency comes at the cost of duplicating individual contributing organizations' local contexts and systems, creating unnecessary – and expensive – redundancies.

A distributed model, on the other hand, meets contributors where they are: organizations that have already invested in archival collection management systems and other locally-hosted methods for publishing finding aids can leverage those solutions, while still participating in the aggregation – without the attendant redundancies. Institutions that lack local infrastructure can use the hosting and publication platform services offered by NAFAN. All institutions benefit from a shared search, browse, and display solution.

Contributor onboarding, training, consultation, and support services

NAFAN will maintain a simple application and onboarding process, whereby any US-based archive can elect to share information about its archival holdings through the network. It will provide a suite of training, consultation, and support services to new and continuing contributors including periodic online trainings, consultation sessions, and a web-based customer service center.

Contributor Dashboard, utilities, and reporting tools

NAFAN will provide a variety of mechanisms to help contributors manage and assess data shared via the network. Examples include:

- A web-based Contributor Dashboard for both contributors and NAFAN staff to manage user accounts, repository profile records, and finding aid sharing workflows on a "self-serve" basis.
- Finding aid data validation and web accessibility checking, with associated reports for contributors. The network could also explore additional data quality reports
- Status and/or any error reports associated with finding aid ingest workflows.
- Web analytics reports summarizing key metrics for repository profile record and/or finding aid usage, and reports summarizing the extent and number of shared finding aids.

Contributor outreach, networking, and resourcing advocacy and investment programs

NAFAN will include an ongoing outreach and marketing program to promote the service to archives and researchers, and will cultivate regional- and peer-based networks to facilitate cross-sharing of information and scale contributor support. In addition to supporting an "ability-to-pay" membership model (thus enabling organizations with little or no funds to have access to tools, workflows, and display solutions), NAFAN will be a strong advocate for grant funding on behalf of archives, and will directly re-invest funds to organizations: as an example, a portion of NAFAN funds (over-and-above that needed for operations and development costs) will be directed back to under-resourced organizations.

6. Next phases

The archives community has an imperative to provide richer and more comprehensive access to collections held by a diverse array of US-based archives – documenting a far broader and representative set of communities, voices, and topics than are currently discoverable. And we can do this in a more sustainable way than we have managed in the past, by abandoning siloed, duplicative infrastructure and leveraging a large-scale, national finding aid network that is community-driven, -sustained, and -governed. This is a timely and overdue opportunity for us to coalesce to meet shared challenges. NAFAN could have a transformative and lasting impact on archives and the researchers they serve by reducing barriers to discovery, expanding the historical record, and offering a new path to sustainable aggregation for archives nationwide.

The findings from the project's research and community workshops make clear the necessity of a national-level network to serve as a discovery service for a comprehensive array of repository profiles and guides to collections stewarded by organizations throughout the US, thus making archival collections more visible and findable on the open web, surfacing connections between collections, and providing researchers with pathways to access those collections. The value of such a network is fundamentally tied to the breadth of its representation of US-based archives: expansive and inclusive participation, representation, sustainability, and governance must be at the core of this effort.

In addition, the network can provide specific benefits to individual organizations by:

- **Enhancing the value of an individual archive's holdings** by situating their repository profiles and finding aids within a national network of related finding aids from academic, community-based, government, historical societies, and public archives – complementing each institution's holdings and amplifying their research value.
- **Supporting archives with determining the level and degree to which they would like to participate and share information about their holdings**, based on their capacity, current workflows, and preferences for the extent of information they choose to provide.
- **Facilitating seamless sharing of finding aids in existing formats** not limited to EAD, with connectors to platforms such as ArchivesSpace, Access to Memory, PastPerfect, and more.
- **Enabling technical capacity for archives while saving cost**, by offering a long-term hosting and publication platform for institutions lacking local infrastructure.
- **Cultivating broader capacity for archives** by providing support, training, tools, and funding resources for archives through a peer-based network.

This report recommends a continued phased, incremental, and community-driven approach that moves this effort from the research project – Phase 1 – to a viable program over a 4-5 year horizon:

Phase 2: Establishing a coalition; expanding project leadership; identifying potential long-term administrative/service hosts, and funding partners (~1-2 years)

Up to this point, CDL has largely led and coordinated the NAFAN project. This next phase will focus on broadening leadership and engagement, as well as securing resources, to advance the project through its next phases of activity. CDL is committed to facilitating this grant-funded phase of work.

- **Establish a coalition of partners** to ensure sufficient community engagement and ownership for Phases 2 and 3 design, planning, and implementation activities, supported by an open framework for participation by archives, researchers, and other stakeholders. It should reflect and be representative of the range of types and sizes of archives in the US, including non-academic research library archives, with funding to support the participation of partners who are not in a position to make in-kind contributions (see [Section 6.2](#)). The coalition will develop a shared leadership model for the project and be guided by a clearly articulated set of values and principles.²¹
- **Expand next-phase project leadership and long-term administrative/service hosts for the network**, to support the coordination of Phase 3 and Phase 4 activities, and conduct follow-up market research to further gauge and hone membership tiers.
- **Identify funding partners and obtain grants** to support Phase 3 and Phase 4 activities.

Phase 3: Developing and launching a "Minimum Viable Product" (MVP) version of the network; garnering community support (~2 years)

This phase would focus on instantiating a robust MVP, functioning as a pilot to demonstrate the value of the network to the broader archives community, enabling them to evaluate the system and make future financial contribution commitments (via the community-based membership program). Any further build-out of more extensive features, still largely grant-funded, will be informed by feedback from this community.

- **Develop the user experience design** (layout, graphic elements, navigation, etc.) and system specifications for the build out of the MVP.
- **Solidify the "ability-to-pay" membership model**, based on up-to-date research regarding archives-specific resourcing levels.
- **Build and launch the MVP** version of the network.
- **Implement a pilot** to populate and test the MVP with data from early participants and get feedback from the community.
- **Conduct a marketing and communications campaign** to promote the network to multiple sectors (funding partners, archives pledging to participate in the proposed membership model, researchers).

²¹ This could begin by revisiting the values and principles established from the NAFAN research project: <https://ucopedu.atlassian.net/wiki/spaces/NAFAN/pages/47744025/Our+Goals+Values+and+Principles>

- **Secure membership pledges from early adopters**, demonstrating community interest and sufficient support for sustainable provision of the NAFAN service once grant funding cycles have ended.

Phase 4: Expanding MVP functionality to support a more fully-featured network and governance structure (~1 years)

This phase is contingent upon securing sufficient up-front membership pledges from the community in Phase 3 to sustain the network moving forward. Assuming this level of financial commitment has been achieved, Phase 4 (the last grant cycle) will focus on building out the remaining prioritized network features, deploying the community-based membership model, supporting archives in migrating to the network, and establishing the governance structure.

6.2 Building the network together

As a community-driven initiative, NAFAN will provide many opportunities to engage in design, planning, and implementation activities across the next three phases of the project. In seeking funding for this initiative, it will be essential to include compensation for partners who are volunteering their time and expertise, as well as partners that would like to participate but are not in a position to make in-kind contributions. Opportunities to participate include, but are not limited to, the following:

- **Planning, design, and development partners:** Partners interested in leading or supporting planning, design, and development initiatives associated with all next phase activities. This includes a range of opportunities, such as:
 - Providing input on and refining initial community-based membership and governance models.
 - Collaborating on user experience design and system specification activities.
 - Conducting or supporting additional research activities, to further hone requirements for the network.
 - Co-promoting and supporting outreach and engagement activities.
 - Providing other in-kind support or resources to help subsidize project activities.
- **Implementation partners:** Partners interested in providing early input on and testing and utilizing the network.
- **Potential long-term administrative/service hosts for the network:** Partners interested in supporting the coordination of Phase 2 and Phase 3 activities (above).
- **Funding partners:** Partners interested in supporting the development of an MVP pilot system, subsidizing the membership model in its initial stage, and helping to launch the fully-featured system.

6.3 Conclusion

The archives community has an imperative to provide richer and more comprehensive access to collections held by a diverse array of US-based archives – documenting a far broader and representative set of communities, voices, and topics than are currently discoverable. And we can do this in a more sustainable way than we have managed in the past, by abandoning siloed, duplicative infrastructure and leveraging a large-scale, national finding aid network that positions us to work together to meet shared challenges and achieve shared goals. It is inadequate to "scale up" our existing regional networks using the same platforms, workflows, and associated service models simply broadened to a much larger contributor base; instead, we need a more expansive set of solutions that address the needs of researchers and archives, as surfaced through the findings presented here, more holistically. NAFAN could have a transformative and lasting impact on archives and the researchers they serve by reducing barriers to discovery, expanding the historical record, and offering a new path to sustainability for archives nationwide.

7. Acknowledgements

The NAFAN research 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. To learn more, visit www.ims.gov.

The project team is indebted to the many individuals who participated anonymously in the research and community workshop engagements, and shared their insight and knowledge with us.

The team thanks those who shared their time and knowledge with us through their participation in the NAFAN research project's Aggregator Partners Working Group, Research Advisory Working Group, and Technical Advisory Working Group. Our work immensely benefited from their perspectives and dedication to this project: Maija Anderson, Hillel Arnold, Carla Arton, Alison Bridger, Jessica BrodeFrank, Celia Caust-Ellenbogen, Aaron Choate, Mark Custer, Kira Dietz, Samantha Dodd, Alexandra Dolan-Mescal, Elizabeth Dunham, Genna Duplisea, Alex Duryee, Lynn Eaton, Karen Eberhart, Maristella Feustle, Kate Flynn, Brenda Gunn, Stacy Haponik, Margaret Kidd, Lee Konrad, Mairelys Lemus-Rojas, Jeanne Long, Steve Majewski, María Matienzo, Holly Mengel, Laurie Lee Moses, Derek Mosley, John Nemmers, Jen Palmentiero, Molly Bruce Patterson, David Pavelich, Kelly Phillips, Ricky Punzalan, Sean Quimby, John Rees, Rebecca Romanchuk, Brian Sheppard, Brian Stevens, William Stingone, Lydia Tang, Marcia Walker-McWilliams, Rachel Walton, Trevor Watkins, and Robert Weaver. The team also thanks the following individuals, who graciously shared information about regional and national aggregation systems implemented in non-US contexts: Kelli Babcock (Archeion); Yves Lapointe (ARCHIVESCANA.ca); Ben Crabstick, Carly Richardson, and Jane Stevenson (Archives Hub); Kerstin Arnold (Archives Portal Europe); Gerhard Müller (Kalliope); Allison Dellit, Julia Hickie, and Simon Jacob (Trove).

CDL appreciates all of the members of the NAFAN research project team, representing multiple partner organizations – Chain Bridge Group, OCLC Research, Shift Collective, and University of Virginia Library – as well as our technical and market research consulting partners, for their expertise and commitment to realizing this project: Keith Bossert, Brittany Brannon, Itza Carbajal, Lori Carlin, Lynn Silipigni Connaway, Raym Crow, Chris Cyr, Christine Di Bella, Brooke Doyle, Erin Glasco, Brian Hoffman, Lynette Johnson, Bergis Jules, Lesley A. Langa, Janet Mason, Daniel Pitti, Merrilee Proffitt, Nicole Sherlock, Zsolt Silberer, Jon Voss, Bruce Washburn, and Chela Scott Weber.

Many CDL colleagues provided input and support in the preparation of this report. Thanks especially to Rachael Hu, Christine Kim, Chad Nelson, and Esther Verreau for their contributions.

Appendix A: Requirements for the network

Overview

Appendix A provides the detailed set of functional requirements and data models necessary for the network. These were developed from the project's research activities²², which included work to:

- understand researcher and archives needs regarding aggregation;
- assess existing aggregation systems and platforms relative to the above identified needs;
- evaluate extent archival description data held by aggregators; and
- identify the relevant current standards and best practices for archival data.

While all of the requirements below are needed for the fully realized network, some are essential to the MVP instantiation (Phase 3), others will be addressed as part of building out the fully-featured network (Phase 4), others will straddle both phases, and others can – and sometimes must – await the realization of a fully-featured network. Each requirement is described below, including whether it will be addressed in either Phase 3 and/or Phase 4, or after, in future releases (see [Section 6](#) for details regarding each phase).

Functions

1. Provide training, consultation, outreach, and tools to cultivate the participation of archives

NAFAN will maintain a simple application and onboarding process, whereby any US-based archive can elect to share information about its archival holdings through the network. The network will provide tools and reports for contributors, to facilitate the management of their repository profile records and/or finding aids. NAFAN will also maintain an ongoing support, outreach, engagement, and advocacy program for contributors, and will also provide support for researcher inquiries.

Table 1.1 Contributor and researcher support services, and outreach

	Requirement	Project phase
1.1	Helpdesk NAFAN will support an online support center ("helpdesk") for contributors and researchers, where individuals can submit questions and requests (e.g., new accounts, technical support, takedown or remediation, report issues (e.g., accessibility issues, errors in finding aids, UI bugs), and access a knowledge base of tutorials and help documentation (e.g., guides for using ancillary tools and reports provided by NAFAN).	Phase 4
1.2	Orientation, training, and ongoing support NAFAN staff will host regularly-scheduled online sessions for contributors at different stages in their engagement with NAFAN including: <ul style="list-style-type: none">• Orientation sessions for potential new contributing organizations;	Phase 3: basic onboarding and support

²² See the complete set of research and community workshop findings at <https://ucopedu.atlassian.net/wiki/spaces/NAFAN/pages/47743740/Project+Reports+and+Resources>

	Requirement	Project phase
	<ul style="list-style-type: none"> • Training sessions covering strategies for sharing repository profile records and finding aids; • Consultation services as needed, to support contributors using NAFAN services. 	Phase 4: remaining requirements
1.3	<p>Contributor networking program</p> <p>NAFAN staff will cultivate regional- and peer-based networks, to facilitate cross-sharing of information and scale contributor support. Examples of potential initiatives include: a peer-based mentoring program; reference and referral coordination to provide support services helping institutions process, describe, and manage collections; hosting online, regional, and/or national meetings. <i>Exemplars:</i> ArchivesSpace member match program; ArchivesSpace member forums</p>	Post-Phase 4
1.4	<p>Marketing and outreach program</p> <p>NAFAN staff will support an ongoing outreach and marketing program (e.g., through newsletters and email communication campaigns) to promote the service to contributors and researchers.</p>	Phase 3

Table 1.2. Contributor onboarding and accounts

	Requirement	Project phase
1.5	<p>Contributor application and new account activation process</p> <p>NAFAN will support a simple application process for archives and regional aggregators to indicate their interest in sharing repository profile records and/or finding aids through the network.</p>	Phase 3
1.6	<p>Contributor user account management</p> <p>NAFAN will maintain user accounts for established contributors, with each account associated with one or more repository profile records. Accounts can be used to log into a Contributor Dashboard.</p>	Phase 3

Table 1.3. Contributor utilities for managing accounts, repository profile records and/or finding aids

	Requirement	Project phase
1.7	<p>Contributor Dashboard</p> <p>NAFAN will support a web-based "self-serve" dashboard for both contributors and NAFAN staff to manage contributor user accounts, repository profile records, and finding aid sharing and redistribution workflows. The dashboard will also include links to contributor reports and ancillary tools. Access to the dashboard will be restricted to individuals with contributor user accounts.</p>	Phase 3
1.8	<p>Ancillary tools</p> <p>NAFAN will collocate and share openly available, third-party ancillary tools to help contributors to create and manage information shared with NAFAN (e.g., validation tools for EAD, MARC, OAI-PMH and other formats; web accessibility</p>	Phase 4

	Requirement	Project phase
	assessment tools; etc.). NAFAN will also provide templates for creating web-accessible collection descriptions in PDF.	

Table 1.4. Contributor reports

	Requirement	Project phase
1.9	<p>Web analytics reports</p> <p>NAFAN will provide contributors with web analytics reports, summarizing key metrics for repository profile record and/or finding aid usage (page views, clickthroughs, sessions, referrals, etc.) for a given timeframe (e.g., annual basis).</p>	<p>Phase 3: basic reports</p> <p>Phase 4: remaining requirements</p>
1.10	<p>Finding aid publication workflow reports</p> <p>NAFAN will provide contributors with essential reports, including:</p> <ul style="list-style-type: none"> • Status and/or error reports associated with workflows for sharing finding aids (including details to facilitate troubleshooting, such as publication state timestamp of finding aid submissions, etc.). • Validation reports for finding aids in structured formats (e.g., XML well-formedness and schema compliance). 	<p>Phase 3: basic reports</p> <p>Phase 4: remaining requirements</p>
1.11	<p>Summary reports</p> <p>NAFAN will provide contributors with reports summarizing the extent and profile of contributed finding aids, for a given timeframe, such as:</p> <ul style="list-style-type: none"> • Number of repository profile records and finding aids, and contribution activity for the year; extent/size of finding aids. • Finding aid data element summaries (e.g., missing descriptive elements, assigned subject/topic headings). • Web accessibility audit reports for finding aids. <p>Post-MVP reports could include potentially harmful or offensive language in repository profile records and/or finding aids.</p>	<p>Phase 3: basic reports</p> <p>Phase 4: remaining requirements</p>

Table 1.5. Contributor advocacy

	Requirement	Project phase
1.12	<p>Contributor investment program</p> <p>NAFAN will invest in building capacity within under-resourced archives: a portion of NAFAN funds (after meeting operations and development costs) will be directed back to under-resourced archives. Potential models include: direct distribution of funds to institutions to support the stewardship of their archival collections (e.g., through micro-grants); and</p>	<p>Post-phase 4</p>

	Requirement	Project phase
	<p>brokering discounts/subsidizing costs for institutions to utilize systems to support the stewardship of their collections (e.g., ArchivesSpace implementation and hosting).</p> <p><i>Exemplar: Mutual Aid at Work initiative: guides, models, and case studies</i></p>	
1.13	<p>Contributor resourcing advocacy</p> <p>NAFAN staff will advocate for processing, description, and conservation/preservation grants, and will lobby for federal/state and private foundation funding on behalf of archives. Beyond funding, NAFAN could advocate for supportive services meeting additional needs, for example: "... a national group to handle requests from the media about relevant materials in small collections, especially during periods of great attention and demand."²³</p>	Phase 4
1.14	<p>Community-based archives consortium advocacy</p> <p>With leadership rooted in CBA's and additional partners, NAFAN could advocate for a consortium of CBAs for mutual aid, support, and collaboration. This includes advocating for a "small institution" no-cost membership category within existing archival organizations such as SAA.</p> <p><i>Exemplar: Community Archives Collaborative</i></p>	Phase 4

2. Support a Registry of repository profiles

NAFAN will maintain a publicly-accessible Registry of archives wishing to share repository profiles and/or finding aids through the network. A given repository profile record will be associated with 1) individual contributor accounts and 2) all collection descriptions contributed by the repository. The [repository profile data model](#) contains descriptive and operational information and is described more fully later in this document. The Registry could be seeded with repository profile records gathered from the IMLS-funded [PROTECCT-GLAM](#) project for contributors to review, edit, and publish as desired.

Table 2.1. Repository profile record management

	Requirement	Project phase
2.1	<p>Repository profile record add/edit utility</p> <p>NAFAN contributors will be able to add and/or edit the descriptive information in their repository profile records. Contributors will co-manage operational information (with NAFAN staff) pertaining to their institution's preferred finding aid sharing workflows.</p>	Phase 3
2.2	<p>Publication process</p> <p>NAFAN will support contributors with "self-serve" options for enabling publication of their repository profile records with or without associated finding aids.</p>	Phase 3

²³ *Increasing discovery of archives: A project to provide better pathways to archival records in cultural heritage collections*, p. 38.

3. Support the sharing of finding aids

NAFAN will support a range of finding aid formats and contribution workflows for both externally and NAFAN hosted content, so that it can be comprehensively aggregated for indexing and display within the Research Portal and optimized to be as "friction-free" as possible, with clear communication regarding processing problems. See the [finding aid data model](#), later in this document.

Table 3.1. Finding aid formats

	Requirement	Project phase
3.1	<p>Accepted finding aid formats</p> <p>NAFAN will ingest finding aids in formats ranging from EAD 2002/EAD3²⁴, MARC21, MARCXML, HTML+Schema.org markup, and PDF (with or without text layers), mining headers, manifest files, and sitemaps to extract structured data.²⁵ Additional unstructured formats such as spreadsheets and documents could be converted to PDF, by contributors and/or by NAFAN as part of the ingestion workflows.</p>	<p>Phase 3: prioritized subset</p> <p>Phase 4: remaining requirements</p>

Table 3.2 Externally-hosted finding aid sharing workflows

	Requirement	Project phase
3.2	<p>Sharing of finding aids through harvesting, web crawling, and fetching processes</p> <p>NAFAN will enable contributors to share finding aids from locally-hosted publication platforms (which serve as their preferred and primary presentation solution for their finding aids), for aggregation within a centralized index and discovery in the NAFAN Research Portal. Ingest mechanisms could include harvesting, web crawling, and/or fetching approaches leveraging OAI-PMH endpoints, sitemaps, manifest files, and APIs (including custom plug-ins or ingest methods with archival collection management systems such as ArchivesSpace, Access to Memory, and PastPerfect). Contributors will co-manage operational information (with NAFAN staff) pertaining to preferred finding aid contribution workflows, via the Registry.</p> <p><i>Exemplars: ArchiveGrid, DigitalNZ, Trove, History of Medicine Finding Aids Consortium²⁶</i></p>	<p>Phase 3: prioritized subset</p> <p>Phase 4: remaining requirements</p>

²⁴ Newer versions of encoding standards should also be supported (e.g., EAD4), as they are implemented across the archives community.

²⁵ The approach of extracting embedded archival descriptions in PDF headers has been utilized by ArchiveGrid (<https://researchworks.oclc.org/archivegrid/about/>); see also Boston College Libraries' use case: <https://dlib.bc.edu/islandora/object/bc-ir:107137>.

²⁶ The National Library of Medicine's (NLM) [History of Medicine Finding Aids Consortium](#) employed a number of these techniques in a distributed hosting model; as of 2022, however, the service was discontinued by NLM.

Table 3.3 NAFAN-hosted finding aid sharing workflows

	Requirement	Project phase
3.3	<p>Sharing of finding aids using a web-based finding aid creation/editing form NAFAN will support a simple web-based utility for creating, editing, and publishing finding aids, conformant with DACS' "Single-Level Optimum" recommendation, on a "self serve" basis, with the option to attach a supplemental PDF version of the guide. <i>Exemplar: Online Archive of California's (OAC) RecordEXPRESS</i></p>	Phase 3
3.4	<p>Sharing of finding aids via uploading process NAFAN will provide single file and bulk upload processes for EAD, MARC, and PDF based finding aids.</p>	Phase 3

Table 3.4 Finding aid preview and publication workflow

	Requirement	Project phase
3.5	<p>Preview and publication process NAFAN will provide a mechanism for contributors and NAFAN staff to preview finding aids in a staging display system, before contributors approve them for publication.</p>	Phase 3

4. Support the discovery and display of aggregated repository profile records and finding aids

NAFAN will support full-text searching, browsing, and display of all aggregated repository profile records and finding aids, within the context of the NAFAN Research Portal. It will promote indexing of the content in commonly-used search engines. It will also support an API to programmatically share content (designated by contributors for third-party sharing and re-use). Researchers will be able to traverse from search and browse results in the Research Portal to a landing page for each contributing organization and each finding aid. The associated [data models](#) are described later in this document.

Table 4.1. Hosting

	Requirement	Project phase
4.1	<p>Collecting guide hosting and publication platform NAFAN will support contributors with a finding aid hosting and publication platform, for contributors choosing NAFAN as their preferred and primary presentation solution for their finding aids, and for aggregation within a centralized index and discovery in the NAFAN Research Portal.</p>	Phase 3
4.2	<p>Persistent URLs NAFAN will assign persistent URLs to all repository profile records and hosted finding aids published in the NAFAN Research Portal.²⁷</p>	Phase 3

²⁷ Content removed per a contributor request will resolve to a 404 page explaining to researchers that the record is no longer available.

Table 4.2. Search

	Requirement	Project phase
4.3	<p>Repository profile record search</p> <p>The NAFAN Research Portal will provide full-text and fielded searching across repository profile records, including support for multilingual content, with facetable search results by holdings topical emphases, geographic location, etc. (based on data as supplied directly by contributors). Post-MVP search functions could employ topic modeling and other techniques to collocate structured and unstructured data for faceting options.</p> <p><i>Exemplar: Archives Portal Europe</i></p>	<p>Phase 3: Prioritized subset</p> <p>Phase 4: remaining requirements</p>
4.4	<p>Finding aid search</p> <p>The NAFAN Research Portal will provide full-text and fielded searching across all finding aids, including support for multilingual content, with facetable search results by dates, creators/contributors, access points, languages, inclusion of links to digital collections, etc. (based on data as supplied directly by contributors). Post-MVP search functions could employ topic modeling and other techniques to collocate structured and unstructured data for faceting options.</p>	<p>Phase 3: Prioritized subset</p> <p>Phase 4: remaining requirements</p>

Table 4.3. Browse

	Requirement	Project phase
4.5	<p>Repository profile record browse</p> <p>The NAFAN Research Portal will support browsing and faceting of repository profile records by standard fields, e.g., organization name, holdings topical emphases, geographic location, etc. (based on data as supplied directly by contributors). Post-MVP browse functions could employ topic modeling and other techniques to collocate structured and unstructured data for faceting options.</p> <p><i>Exemplars: Archives Portal Europe, LA as Subject</i></p>	<p>Phase 3: Prioritized subset</p> <p>Phase 4: remaining requirements</p>
4.6	<p>Finding aid browse</p> <p>The NAFAN Research Portal will support browsing and faceting of finding aids by standard fields: dates, creators/contributors, access points, languages, inclusion of links to digital collections, etc. (based on data as supplied directly by contributors). Post-MVP browse functions could employ topic modeling and other techniques to collocate structured and unstructured data for faceting options.</p>	<p>Phase 3: Prioritized subset</p> <p>Phase 4: remaining requirements</p>

Table 4.4. Repository profile record and finding aid display

	Requirement	Project phase
4.7.	<p>Repository profile record landing page</p> <p>The NAFAN Research Portal will support a landing page/view for each repository profile record with a persistent URL maintained by NAFAN.</p> <p><i>Exemplars: Archives Portal Europe, LA as Subject</i></p>	Phase 3

	Requirement	Project phase
4.8.	<p>Finding aid landing page</p> <p>The NAFAN Research Portal will support a landing page/view with a NAFAN maintained persistent URL for each NAFAN-hosted finding aid. Landing pages to externally-hosted finding aids will include a link to view the full finding aid within the context of the contributing organization's locally-hosted publication platform. Landing pages for NAFAN-hosted finding aids will provide researchers with the option to view the full finding aid within the Research Portal.</p> <p><i>Exemplar: ArchiveGrid (for displays of externally hosted finding aids)</i></p>	Phase 3
4.9.	<p>NAFAN hosted finding aid detailed view</p> <p>The NAFAN Research Portal will support a detailed view for NAFAN-hosted finding aids, including additional features such as the ability to search within that guide and links to contact the institution to submit reference and photoduplication requests.²⁸ The display functions could include options for archives to specify descriptions that require mediated access for researchers to view (e.g., via user account-based logins).</p>	Phase 3

Table 4.5. Search engine optimization and data re-use

	Requirement	Project phase
4.10	<p>Search engine optimization and indexing</p> <p>NAFAN will promote indexing of the aggregated repository profile records and finding aids in commonly-used search engines.</p>	Phase 3
4.11	<p>Application Programming Interface (API)</p> <p>NAFAN will support a publicly-documented API for programmatically accessing repository profile records and finding aids, as allowed and specified by contributors' based on "opt-in" decisions.</p>	Phase 4

5. General service requirements

NAFAN will maintain a robust operations and underlying information technology infrastructure based on industry standards for continuity and security. Practices will include maintaining appropriate and up to date policies, system resource monitoring, high availability configuration, regular auditing and operating system updates and change management controls, and intrusion detection and electronic access controls in three distinct environments: development, testing, and production.

Table 5.1. Service policies and operations

	Requirement	Project phase
5.1	<p>Administration and policy documentation</p>	Phase 3: Prioritized subset

²⁸ This includes links to relay photoduplication and holdings requests to external systems implemented by contributors, such as [Aeon](#).

	Requirement	Project phase
	<p>NAFAN will maintain and provide publicly-accessible versions of all policy documentation, for reference by contributors and researchers. Example policies include:</p> <ul style="list-style-type: none"> ● Terms of service agreement ● Contributor eligibility and participation requirements ● Finding aid requirements ● Terms for sharing and re-use of contributors' repository profile records and finding aids ● Privacy/removal of user data ● Web accessibility compliance and recommendations for remediating non-web accessible collection ● Removal and takedown procedures ● Service level agreements, including with third parties ● Data recovery/backup procedures 	Phase 4: remaining requirements
5.2	<p>Web accessibility compliance All NAFAN user interfaces will comply with relevant Web Content Accessibility Guidelines (WCAG) standards (or Section 508, if NAFAN is maintained by a federal agency). NAFAN will support protocols for identifying and generating reports of finding aids that are not web accessible, will provide a mechanism for users to report content that is not accessible, and will confer with contributors on strategies, tools, and best practices to remediate the finding aids.</p>	Phase 3
5.3	<p>Privacy compliance Reports will comply with community-based, national, and international privacy statutes and regulations, such as the General Data Protection Regulation (GDPR).</p>	Phase 3
5.4	<p>NAFAN administrator accounts NAFAN staff will have full administrative level permissions to configure system functions, manage repository and finding aid records, create and manage contributor accounts and permissions, impersonate users, and access contributor reports.</p>	Phase 3

Data models

NAFAN's search, browse, and display functionality hinge on robust data models for both repository profiles and finding aids, both of which are detailed below. Contributor supplied profile data will be provided and managed via the Registry. As included in the element descriptions, the models will conform to the [International Standard for Archival Description \(General \(ISAD-G\)\)](#) and [DACS' "Single-Level Optimum"](#) recommendations.

Repository profile record data model

The repository profile data model will contain descriptive and operational information:

- **Descriptive information** supplied and maintained by the contributor, for indexing and display in the NAFAN Research Portal.
- **Operational information** used internally by NAFAN staff for administrative purposes, and to assist contributors with sharing and maintaining their repository profile records and finding aids.

Table 6. Descriptive information

Element	Description	Data source	Status
Unique identifier	An identifier to uniquely distinguish each contributing organization represented in the Registry. Conforms to ISDIAH 5.1.1 Identifier. Any additional identifiers, related to entries in external services (e.g., SNAC, Research Organization Registry, Virtual International Authority File, Wikidata, etc.), could also be managed in the Registry.	NAFAN derived and assigned.	Required
Persistent URL	A persistent URL for each repository profile record published in the NAFAN Research Portal.	NAFAN derived and assigned.	Required
Repository name	A human readable name to be indexed and displayed within the context of repository profile records and/or finding aids in the NAFAN Research Portal. Conforms to ISDIAH 5.1.2 Authorized form(s) of name.	Contributor supplied via Dashboard.	Required
Repository type	Identifies the contributing organization type. Conforms to ISDIAH 5.1.5 Type of institution with archival holdings.	NAFAN derived and assigned; editable by contributor via Dashboard.	Required
Location and address(es)	All relevant addresses for the contributing organization, both physical and electronic. Conforms to ISDIAH 5.2.1 Location and	Contributor supplied via Dashboard.	Required

Element	Description	Data source	Status
	address(es).		
Contact information	Details necessary for researchers to contact the contributing organization (e-mail, phone, etc.). Conforms to ISDIAH 5.2.2 Telephone, fax and email.	Contributor supplied via Dashboard.	Required
Parent organization name	A parent-level organization, encompassing multiple individual contributing organizations (e.g., a university, county library system, etc.). Also used to group individual contributing organizations related to the same parent organization in browse results.	Contributor supplied via Dashboard.	Optional
Latitude / longitude	Geo-coordinates for the location of the contributing organization.	NAFAN derived and assigned; editable by contributor via Dashboard.	Optional
Holdings description	A profile of the holdings of the contributing organization, including any access restrictions. Conforms to ISDIAH 5.3.7 Archival and other holdings.	Contributor supplied via Dashboard.	Optional
Holdings topical emphases	Topical subject headings (strings or URIs), characterizing key collection scope(s) and emphases of the contributing organization and, as exists, the source for the term (e.g., the thesaurus, URL, and/or descriptive rules or guidelines used to create the term).	Contributor supplied via Dashboard.	Optional
Digital collections	URL to the contributing organization's digital collections site(s)/platform(s).	Contributor supplied via Dashboard.	Optional
Photoduplication / public service reference request systems	URL to the contributing organization's preferred photoduplication / public service reference request systems (e.g., Aeon) to direct researcher inquiries.	Contributor supplied via Dashboard.	Optional

Table 7. Operational information

Element	Description	Data source	Status
Finding aid harvesting specifications	URL endpoint(s) and protocol details, to facilitate NAFAN harvesting, web crawling, or fetching of finding aids shared by the contributing organization. Also includes a specified frequency for re-harvesting.	NAFAN specified, in consultation with the contributor.	Required
Repository profile and finding aid redistribution preferences	Contributing organizations can opt to specify if they would like their repository profiles and/or finding aids in NAFAN redistributed with third-parties, through the NAFAN API or other protocols (e.g., bulk download, OAI-PMH data provider service).	Contributor supplied via Dashboard.	Optional
Finding aid default data values	Contributing organizations can opt to specify default values for their finding aids shared with NAFAN. (See the complete proposed finding aid data model below).	NAFAN specified, in consultation with the contributor.	Optional

Finding aid data model

Detailed below are the minimum set of structured data elements that will be requested from contributing organizations to support baseline indexing and display of their finding aids within the NAFAN Research Portal. NAFAN will extract or derive a structured descriptive record for the collection from the encoded content itself (see [Requirement 3.1, Accepted finding aid formats](#)).

NAFAN will provide contributors with documentation and tutorials outlining mappings used to derive the records, and will host periodic virtual training sessions to highlight strategies for providing structured data when sharing finding aids. If the requested data elements are not available in the source collection description, NAFAN will supply a default data value (e.g., "This information is not available in the source collection description; please contact the contributing organization for more information").

Table 8. Finding aid information

Element	Description	Data source	Status
Unique identifier	A unique identifier for the finding aid. NAFAN will leverage the file name for the finding aid, to serve as a unique identifier.	Extracted from submitted finding aid.	Required
Persistent URL	A persistent URL for each finding aid hosted by NAFAN.	NAFAN assigned.	Required
Reference code	A unique identifier for the collection. Conforms to ISAD(G) 3.1.1 Reference code and DACS 2.1 Reference code element.	Extracted from submitted finding aid.	Required

Element	Description	Data source	Status
Name and location of repository(s)	Name and location of the contributing organization(s) stewarding the collection. Conforms to ISDIAH 5.1.2 Authorized form(s) of name, ISDIAH 5.2.1 Location and address(es), and DACS 2.2 Name and location of repository.	Contributor supplied via the Dashboard (NAFAN will apply the data to all finding aids)	Required
Title	A title for the collection. Conforms to ISAD(G) 3.1.2 Title and DACS 2.3 Title.	Extracted from submitted finding aid.	Required
Level of descriptions	Level of description for the collection (e.g., fonds, collection, series, subseries, etc.). Conforms to ISAD(G) 3.1.4 Level of description and DACS 1 Levels of description.	Extracted from submitted finding aid.	Optional
Dates	Date(s) associated with the creation of the collection. Conforms to ISAD(G) 3.1.3 Date(s) and DACS 2.4 Date.	Extracted from submitted finding aid.	Optional
Extent	Extent and physical nature of the collection. Conforms to ISAD(G) 3.2.1 Extent and medium of the unit of description (quantity, bulk, or size) and DACS 2.5 Extent.	Extracted from submitted finding aid.	Optional
Creators / contributors	Corporate bodies, persons, and families associated with creating and/or contributing to the collection. Conforms to ISAD(G) 3.2.1 Name of creator(s) and DACS 2.6 Name of creator(s).	Extracted from submitted finding aid.	Optional
Scope and content	Description of the materials and activities reflected in the collection. Conforms to ISAD(G) 3.3.1 Scope and content and DACS 3.1 Scope and content.	Extracted from submitted finding aid.	Optional
Conditions governing access	Access restrictions related to the collection. Conforms to ISAD(G) 3.4.1 Conditions governing access and DACS 4.1 Conditions governing access.	Extracted from submitted finding aid; or contributor supplied via the Dashboard (NAFAN will apply the data to all finding aids that do not have this data).	Optional

Element	Description	Data source	Status
Conditions governing use	Use restrictions related to the collection. Conforms to ISAD(G) 3.4.1 Conditions governing reproduction and DACS 4.4 Conditions governing reproduction and use.	Extracted from submitted finding aid; or contributor supplied via the Dashboard (NAFAN will apply the data to all finding aids that do not have this data).	Optional
Languages and scripts of the material	Language(s), script(s), and symbol systems represented in the collection. Conforms to ISAD(G) 3.4.3 Languages/scripts of material and DACS 4.5 Languages and scripts of the material.	Extracted from submitted finding aid.	Optional
Finding aid license / usage notice	Conditions under which the finding aid is available for use and usage notices (e.g., a URI to a Creative Commons license, Local Contexts institution notice, etc.). Conforms to DACS 8.2 Rights statements for archival description.	Extracted from submitted finding aid; or contributor supplied via the Dashboard (NAFAN will apply the data to all finding aids that do not have this data).	Optional
Finding aid creation and revision history	Notes related to the creation of the finding aid, and/or revisions and remediations to the finding aid. Conforms to DACS 8.1 Description control.	Extracted from submitted finding aid.	Optional
Administrative / biographical histories	A biographical or administrative historical description of the creators of the collection. Conforms to ISAD(G) 3.2.2 Administrative/biographical history and DACS 2.7 Administrative/biographical history.	Extracted from submitted finding aid.	Optional
Access points	Controlled indexing terms for names (e.g., corporate bodies, persons, and families), places, subjects, documentary forms, etc. represented in the collection.	Extracted from submitted finding aid.	Optional

Element	Description	Data source	Status
Digital collections	URLs to digital collections site(s)/platform(s) (including sites managed by the contributing organization, state/regional digital collection aggregators), with digital versions of the materials described in the finding aid.	Extracted from submitted finding aid.	Optional
Additional finding aids	URLs for additional versions of the finding aid. Conforms to ISAD(G) 3.4.5 Finding aids and DACS 4.6 Finding aids.	Extracted from submitted finding aid.	Optional

Appendix B: Assessment of existing systems

The following summary reflects the result of an environmental scan and evaluation of potential solutions, using the following rubric:

- **Features:** Solutions that support as many of the requirements outlined in [Appendix A](#) as possible, and that have been developed for the presentation of finding aids as a key use case – with support for standardized structured data formats (e.g., EAD, MARC21, MARCXML, HTML) and ideally, unstructured formats such as PDF.
- **Code:** Solutions that are built using open-source code, and are actively maintained by the implementor communities. While there are attendant costs and considerations with such systems, the benefits – which include more adaptable and flexible system frameworks, supported and sustained by a peer-based community of implementers – outweigh those of commercial solutions.
- **Hosting model:** Solutions that can be readily adapted to support the systems-oriented features and functions for the network in a distributed content model, or solutions that can be selectively adapted to support a specific feature and function of the network (as a modularized component within a distributed hosting model).

The assessment included reviewing publicly-available documentation summarizing the system functions, use cases, and technology stack; and conducting information gathering from organizations that maintain or use these systems. The summaries below also include descriptions of any significant gaps in addressing areas of prioritized need.

Custom systems implemented by other state/regional and national aggregators

This category includes the custom systems most widely developed and implemented by state/regional and national aggregators: [Archives Portal Europe](#), [Kalliope](#), [Philadelphia Area Archives](#), [Rhode Island Archives and Manuscripts \(RIAMCO\)](#), and [Texas Archival Resources Online \(TARO\)](#).

- **Features:** These custom-built solutions support the discovery and display of aggregated repository profile records and finding aids via a research portal, ingest-based finding aid contribution and publication workflows, and a registry of repository profile records. They are optimized for ingesting, indexing, and publishing finding aids in the EAD format. Archives Portal Europe and Kalliope additionally support ingesting, indexing, and publishing of biographical and historical information about persons, families, and organizations that created archival collections (or are documented in archival collections), in the [Encoded Archival Context for Corporate Bodies, Persons, and Families \(EAC-CPF\)](#) format.
- **Code:** The implementor communities are limited to the organizations that developed and are using the code.
- **Hosting model:** These systems are based on a centralized hosting model.

- **Gaps:** These systems would need to be refactored to support additional structured and non-structured finding aid formats, with additional infrastructure components to support a broader range of ingest, web crawling, and harvesting-based finding aid contribution workflows.

Archival management systems: ArchivesSpace and Access to Memory

[ArchivesSpace](#) and [Access to Memory](#) represent the two most predominant open-source and currently maintained archival collection management systems in North America (developed by ArchivesSpace and Artefactual, respectively). [ARCHIVESCANADA.ca](#) and [Archeion / Archives Association of Ontario](#) have implemented Access to Memory as their core aggregation infrastructure, in a centralized hosting model.

- **Features:** As archival collection management systems, these applications are designed to support the complete lifecycle for archival collection management.²⁹ They support the discovery and display of aggregated repository profile records and finding aids via a research portal, finding aid contribution and publication workflows, and a registry of repository profile records. They can ingest finding aids in EAD, MARC, and other structured data formats.
- **Code:** These open source systems are actively supported by their host development organizations, and have active community-based code contributions.
- **Hosting model:** These systems are based on a centralized hosting model.
- **Gaps:** ArchivesSpace and Access to Memory provide a broad range of collection management functions not needed by NAFAN, such as accession management, shelving locations, donor tracking, etc. Within a multi-tenant deployment context, these applications are not designed to accommodate tailoring for individual institutions' description preferences (e.g., data requirements for finding aids). Additional challenges would include isolating institution-specific data (e.g., preferred controlled vocabularies and taxonomies). More generally, they would also require refactoring to support additional non-structured finding aid formats and additional infrastructure components to support a broader range of ingest, web crawling, and harvesting-based finding aid contribution workflows.

Index and display systems optimized for indexing and display of finding aids: ArcLight

[ArcLight](#) is an open-source [Blacklight](#)-based indexing and discovery system designed specifically for archival collections, developed by Stanford Libraries with partner organizations. [Connecticut's Archives Online \(CAO\)](#) and [Empire Archival Discovery Cooperative \(EADC\)](#) are two reference models of ArcLight implementations within state/regional finding aid aggregation contexts.

- **Features:** ArcLight is optimized for the indexing and display of aggregated repository profile records, finding aids, and related digital content via a research portal, and a registry of repository profile records. It is optimized to support indexing, and publishing finding aids in the EAD format. It does not include finding aid contribution and publication workflow support;

²⁹ "Archival management systems typically provide integrated support for the archival workflow, including appraisal, accessioning, description, arrangement, publication of finding aids, collection management, and preservation." Spiro, Lisa. (2009), *Archival Management Software*, https://www.clir.org/wp-content/uploads/sites/6/spiro_Jan13.pdf, p. 2

components for ingesting, crawling, or harvesting finding aids are not part of the ArcLight system.

- **Code:** It is built on the currently maintained open source Blacklight code base with active community-based code contributions.
- **Hosting model:** ArcLight could potentially be adapted as a component within a centralized or distributed hosting model.
- **Gaps:** Because of its focus on EAD finding aids, ArcLight would need to be adapted to support a broader range of finding aid formats, and other components would need to be developed to support the broader functions of the network (e.g., finding aid contribution workflows). Given its primary emphasis on indexing and display, ArcLight is not a strong candidate for a full-service solution, but could be extended to support indexing and display of additional structured and non-structured finding aid formats, and serve as one modularized component within the broader suite of custom-built infrastructure needed for the national network.

Social Networks and Archival Context (SNAC)

[SNAC](#) is a cooperative program and international aggregation of biographical and historical information about persons, families, and organizations that created archival collections (or are documented in archival collections), and their connections to one another. The service is managed by the University of Virginia Library and US National Archives and Records Administration.

- **Features:** SNAC is optimized for the aggregation, indexing, and display of identity records for persons, families, and organizations, highlighting their connections to one another and associations with archival collections. The aggregation also includes repository profile records for individual archives. SNAC's data model for identity records leverages the EAC-CPF standard; the bulk of the identity records were originally derived between 2010-2015 from MARC records exported from WorldCat, and EAD finding aid data gathered from regional aggregators and individual institutions.
- **Code:** SNAC is built on open-source code, with code contributions primarily from the SNAC development team.
- **Hosting model:** The system is based on a centralized hosting model.
- **Gaps:** SNAC's emphasis is on the aggregation of identity records, so the existing technical architecture would need refactoring to support the aggregation of finding aids and associated contribution workflows. Given its specific focus, it could be integrated with NAFAN to support high-value features such as cross-linking of finding aids published through NAFAN with identity records published through SNAC. Additional integrations could be explored over time, and initially facilitated by incorporating SNAC identifiers into NAFAN repository profile records.

ArchiveGrid

[ArchiveGrid](#) is a service managed by OCLC Research. It comprises an international aggregation of over seven million finding aids, and serves as a basis for experimentation and testing in text mining, data analysis, and discovery system applications and interfaces.³⁰

- **Features:** ArchiveGrid supports the discovery and display of aggregated repository profile records and finding aids via a research portal, flexible finding aid contribution and publication workflows (for a range of formats), and a registry of repository profile records. It predominantly comprises MARC records for archival collections, selectively harvested from WorldCat. It supports finding aid data in HTML, EAD, and PDF formats – including ingest, web crawling, and metadata harvesting approaches.
- **Code:** ArchiveGrid is a proprietary system built on open-source code, and maintained by OCLC Research.
- **Hosting model:** The system is based on a distributed hosting model. ArchiveGrid centrally aggregates and indexes finding aids; to view the full finding aid details, researchers are routed from the ArchiveGrid to the endpoint from which the finding aid was originally harvested, web crawled, and/or ingested.
- **Gaps:** While ArchiveGrid supports a range of systems-functions proposed for the network and is a de facto example of the distributed hosting model, it does not support the hosting and publication of finding aids and the underlying code is not readily accessible to more fully adapt to the proposed requirements. It can, however, serve as an exemplar and reference model for designing and implementing ingest, web crawling, and metadata harvesting techniques.³¹

³⁰ See <https://www.oclc.org/research/areas/research-collections/archivegrid.html>

³¹ The National Library of Medicine's (NLM) [History of Medicine Finding Aids Consortium](#) also employed a number of these techniques in a distributed hosting model; as of 2022, however, the service was discontinued by NLM. Digital aggregations such as DigitalNZ and Trove serve as additional and current reference models that utilize these approaches.