

**Society of American Archivists
Council Meeting
November 2-3, 2018
Chicago, Illinois**

**Recommendation to Establish a
Committee on Research, Data, and Assessment**

(Prepared by the Task Force on Research/Data and Evaluation)

BACKGROUND

In May 2017, the SAA Council charged the Task Force on Research/Data and Evaluation (TF-CORDE), comprised of Chair Michelle Light and member Sarah Buchanan, Mahnaz Ghaznavi, Dennis Meissner, Daniel Noonan, and Stacie Williams, to complete the following tasks by November 2018:

- Determine whether sufficient need exists to justify the effort and costs associated with establishing a standing body to conduct, facilitate, and/or evaluate research that is meaningful for SAA and the archival community.
- Determine how such a standing body should be structured, staffed, and governed.
- Determine how such a standing body might be financed and supported.
- Propose how such a standing body might interact with other SAA groups.
- Propose how such a standing body might interact with external groups.

In his SAA presidential address, "[Bare Necessities](#)," Dennis Meissner argued that SAA must "concentrate on gathering, evaluating, and presenting the real quantitative and qualitative evidence that supports all the compelling narratives and theoretical arguments about the value of archives. We need this evidence because we have struggled for many years with the challenge of demonstrating the 'value' of archives via anything resembling objective measures." He proposed that SAA form a Committee on Research and Evaluation (CORE), with a goal "to provide access to compelling data about American archives and their users that speak to the value of archives for society and that also help us improve our services to our consumers." In 2016 he [proposed CORE to the Council](#) and outlined potential areas of responsibilities. As a result, the SAA Council created TF-CORDE to investigate need and propose a charge for such a group.

DISCUSSION

As [reported to the Council in May 2018](#), TF-CORDE conducted interviews and created personas to learn about the data and research practices and needs of SAA members and SAA staff. TF-CORDE identified five needs for:

- 1) Standardized tools for gathering and analyzing data;
- 2) A centralized repository of data, tools, and other authoritative aids;

- 3) Training on gathering, analyzing, interpreting, and using data;
- 4) Up-to-date, basic facts and figures about archives and archivists; and
- 5) A clearinghouse to support archival surveys and research.

The May 2018 report contains explanations for each need. The personas will be posted onto the TF-CORDE microsite.

Between May and November 2018, TF-CORDE next examined how an SAA standing body might assume responsibility for addressing these needs. TF-CORDE examined similar standing bodies within other professional organizations to make these recommendations. Recognizing that we cannot ask SAA staff to assume additional responsibilities, that resources are limited, and that a standing body would depend on volunteer expertise, TF-CORDE explored how SAA might address each of the identified needs.

1. Standardized Tools for Gathering and Analyzing Data

TF-CORDE identified a need for simple, no-cost, standardized survey tools to help archivists gather information about several archival functions (e.g., public service interactions, collection management activities, etc.) at their own institutions, so they could compare their performance to others and/or evaluate their institution's effectiveness over time.

There are several efforts underway that may form a foundation for the development of standardized tools for gathering and analyzing data essential for archivists and their stakeholders. The SAA-ACRL/RBMS Joint Task Force on the Development of Standardized Holdings Counts and Measures for Archival Repositories and Special Collections Libraries is developing standardized measures for quantifying holdings, which should support the aggregation of holdings information across multiple repositories so researchers may be able to report on or estimate how much archival material exists in U.S. repositories. The [Standardized Statistical Measures and Metrics for Public Services in Archival Repositories and Special Collections Libraries](#) was approved as a standard by ACRL in October 2017 and by the SAA Council in January 2018. This standard is designed to “help archival repositories and special collections libraries develop local statistical data collection policies and practices around their public services so that they might in turn use the data to assess their public service operations.”

Between 2009 and 2016, [Archival Metrics](#), a joint project of the University of Michigan, the University of North Carolina-Chapel Hill, and the University of Toronto, funded by the Andrew W. Mellon Foundation and the National Historical Publications & Records Commission (NHPRC), developed standardized evaluation toolkits that included validated, tested questionnaires, administration and coding instructions, and sample reports to help repositories gather and assess data about how effectively they were serving users' needs. Our persona interviews revealed that archivists still need and demand these ready-to-use evaluation toolkits to help them assess their performance longitudinally or benchmark to others; however, they did not reference these resources specifically, and several wished to measure performance in other areas as well, such as processing productivity. Nevertheless, this project could be a resource to promote further work or might provide some valuable lessons about offering such toolkits.

Whereas there are several examples of groups and individual researchers conducting surveys about archives, archivists, and similar cultural heritage organizations and professionals, there are fewer examples of professional organizations who have provided standardized survey tools for regular, local implementation. Aspirational models might include those from [DataArts](#), the [Council for the Advancement of Standards in Higher Education](#), the [Council for Advancement and Support for Education](#), or the [Association for Research Libraries](#). SAA lacks the resources and infrastructure to emulate these robust examples; however, TF-CORDE believes it would be possible for a standing SAA committee, with additional research and consultation with other SAA component groups, to either 1) identify, explain, promote, and potentially enhance one or more survey tools, or 2) develop and promote one or more simple survey tools. A standing SAA committee might develop a resource akin to the Standards Portal for promoting and explaining these tools as well as highlighting resources for how to analyze the data. A longer-term goal for the standing SAA committee would be to identify a sustainable platform for collecting this data from individual institutions. Ideally, this collected data would allow institutions to assess themselves over time, but could also be leveraged for other cross-institutional or profession-wide analysis.

2. Centralized Repository for Data and Other Research Outputs

Centralized Repository

TF-CORDE identified a shared desire for a centralized, online repository where archivists could find and/or contribute data and other resources useful for archives and archivists. Users wanted one consistent place to find, use, and contribute raw data from published research products, data from salary surveys, and data from other smaller surveys, institutional assessments or research. While we recognized that data and other resources do not need to be deposited in a single repository to be discoverable in a single place, we also are sensitive to the fact that many archivists and repositories do not have access to a repository, nor are familiar with a repository that would potentially house their data or resources.

Although, the existing SAA website might provide a centralized portal for this content, both as a means to link out to content that is housed in other repositories and to house content with no other home, the website lacks several important features, such as providing for persistent identifiers or offering preservation functionality. We therefore investigated potential options for hosting data and resources relating to archives, with the expectation that several individuals have used or will use other options, such as their own institutional repositories, ICPSR, OSF, github, or other publishing options.

TF-CORDE investigated the following repository options:

- [ALAIR](#): This is the American Library Association's repository on the D-Space platform. ALA outsources the hosting and management to the University of Illinois Urbana-Champaign Archives. This model seemed too resource intensive.
- Open library and archives repositories: There are some freely available repositories relating to library and archives, but not yet widely used. Some examples include the [LIS Scholarship Archive via OSF Preprints](#) and the library science section of [Humanities](#)

[Commons](#). While these repository platforms allow contributors to share and preserve their data and other formats, they largely just allow users to download files.

- [Dataverse Project](#): This specialized, open-source, web-based data platform supports the analysis of tabular data, in addition to offering sharing, preservation, and citation functions. Individuals and institutions can install and manage their own instances of Dataverse and federate with others to increase discoverability, or use the [Harvard Dataverse](#) to set up their own customized Dataverse.
- [Odum Institute for Research in Social Science](#): Cal Lee, Editor of the *American Archivist*, plans to work with the Odum Institute at UNC Chapel Hill to set up a Dataverse dedicated to housing data referenced in *American Archivist* articles when authors do not have other options for preserving and sharing their data in their own institutional repository. If this comes to fruition, then it makes sense to explore how to expand on this work to meet broader professional needs.¹

Setting up a repository, helping archivists prepare their data properly, reviewing data for suitability, and managing the entire deposit and publication process would obviously take a significant investment of time and expertise. Due to this complexity and expressed breadth of the community's desire for a data repository and portal, TF-CORDE feels it cannot recommend a single solution at this time. We recommend that the standing committee build upon the findings here in scoping and recommending a solution. First steps should include compiling a web-based resource to guide archivists on where to find relevant data, explore further the collaborative possibilities for publishing archives-related data with the *American Archivist*, and developing guidance for how archivists might preserve and share their research data in other trusted repositories until a standing committee identifies a better option.

Repository Attributes

Based on input from the personas project and our repository investigations, the Task Force identified a core set of recommended and desirable features for a repository.

- Recommended:
 - Searchability
 - Download content
 - User submissions
 - Workflow to support deposit agreements and adherence to data best practices²

¹ TF-CORDE discussed how a proposed standing committee might work with the *American Archivist* editor and editorial board to consider broadening the scope of its publication program to accept, review, and publish data sets. Consider a Data Portal with a Data Editor similar to the Reviews Portal and Editor. If data were another stream of research content submitted for publication, then a Data Editor might advise and help contributors on how to prepare their data. A peer review process could evaluate quality and whether the data follow [FAIR best practices](#) so the data are findable, accessible, interoperable, and reusable. The publication process could ensure the data are described, findable, and preserved. Complementary to this model is the [data paper](#), a searchable metadata document that describes a particular dataset or a group of datasets and is published in the form of a peer-reviewed article in a scholarly journal.

² TF-CORDE discussed the importance of further considering deposit requirements and agreements to ensure contributed data serve professional needs. Significantly, TF-CORDE found that archivists may need education and

- Versioning
- Assignment of a persistent identifier or DOI
- Bit-level preservation
- Desirable:
 - Access and user management
 - Tools for interacting with data
 - TRAC features
 - Usage statistics/analytics
 - APIs for integrating into external tools

3. Training on Gathering, Analyzing, Interpreting, and Using Data

While SAA’s Committee on Education has tailored a curriculum that is responsive to the needs of the membership, there currently remains a paucity of professional development opportunities designed to cultivate data gathering, analysis, interpretation and use skills and strategies.

To address this need, a two pronged approach may be useful: in the near and intermediate term, the Committee on Education may identify courses offered by other organizations that support data intensive research skills and strategies while at the same time working with SAA to identify longer term responses, such as identification of available faculty to teach courses sponsored by the association, or partnership opportunities with peer organizations.

4. Collecting/Providing Access to Up-To-Date, Basic Facts About Archives and Archivists

A key finding of the personas exercise is that, at a minimum, SAA members would like to have access to current and longitudinal data regarding:

- Archival repositories - size (physical, staff and holdings) and location;
- Basic demographics about the archival profession including, among other data, collected items regarding diversity and granular salary and benefits information; and
- Archival graduate programs.

To this end, the task force reviewed existing resources that might satisfy these needs. The [Facts & Figures](#) page, one of several grouped under the Resources section of SAA’s public facing web site, functions as a clearinghouse for site visitors to gain quick access to information about the profession. The page features access to reports, survey results, and select published data sets through a combination of outbound links to resources that may live in SAA’s domain or on external sites. The page is a good start. At the same time, the task force recognizes a number of missed opportunities associated with how facts and figures are captured and/or presented to users on this page:

- Lengthy print reports are typically not accompanied by dashboards (or similar key messages) that would enable a quick assimilation of important facts or access to underlying data sets;

assistance in creating and contributing data that would follow [FAIR best practices](#) to ensure data are findable, accessible, interoperable, and reusable.

- Occasional reports by affiliated organizations do not consistently offer cumulative integration of their findings or any distillations and promotion of their key messages for archivists;
- Overarching themes and *key data points* across reports are not readily available; and
- Current and/or longitudinal data about archives, archivist and issues related to the profession are not readily available.

These deficiencies, except for the last bullet point, could be dramatically improved by greatly enlarging the *Facts and Figures* structure to accommodate a much wider range and depth of information, curated and otherwise, to help archivists both better understand their environment and to enhance their performance as professionals. The following table illustrates briefly how such an expanded structure might look and the sorts of informational resources it could deliver to members. It divides potential content into three broad channels and suggests the sort of content—some extant and some potential—with which it could be populated.³ This is simply a suggested structure and not a considered recommendation:

Statistics & Analytics	Project Reports & Tools	Data Sets
<p>Content types: Dashboards, fact sheets, at-a-glance stat tables, and graphs would be the most common presentation formats of these primarily quantitative data. Pre-digested quantitative and qualitative information gathered together from diverse sources. Relevant questions treated might include: Who are American archivists? Who employs them? Whom do they serve and how do they serve them? What economic benefits do they provide? What social and cultural benefits do they provide?</p>	<p>Content types: Formal presentations of research findings and information on and from projects currently underway, as well as sharable tools resulting from these projects. CORDE could commission or support research and collaboration projects aimed at producing reports which could influence standards and best practices across U.S. repositories, or which could produce practical tools that archivists can use to achieve better results in their normal practices. These endeavors could include projects commissioned by SAA CORDE, projects commissioned by SAA committees and component groups, and projects commissioned by external agencies. “Tools” could include not just technical tools (e.g., software, templates), but also</p>	<p>Content types: Underlying data sets from SAA projects, as well as useful data from external sources. These data aggregations could presumably be harvested in whole or in part from the microsite or interrogated via APIs.</p>

³ A fuller version of the [chart](#) contains many examples of specific content (known and speculated) that could populate an expanded *Facts & Figures* page.

	assessment tools to help repositories evaluate their performance against their peers or against accepted standards and best practices.	
--	----------------------------------------------------------------------------------------------------------------------------------------	--

Several similar cultural heritage organizations conduct and maintain annual surveys that gather information regarding salaries and other financial considerations, workforce/profession makeup, diversity, and training needs. This is typically handled by an association staffer, possibly within the membership staff and/or contracted out to a third party. TF-CORDE discussed how a standing committee might identify and prioritize these needs, then advocate or seek annually the necessary resources to complete them.

5. Clearinghouse to Support Archival Surveys and Research

Development of Research Agenda

In early 2018, TF-CORDE determined that one of the most effective ways to capture the long-term research aspirations, communicate them inclusively to SAA members, and connect past research accomplishments such A*CENSUS with future research goals would be to develop a research agenda for the archival profession. Such a document was prepared for our June Task Force meeting and was subsequently refined by Task Members’ input, building on our shared insights from completing the interviews for our Personas. TF-CORDE displayed our Research Agenda in poster format at the 2018 SAA Annual Meeting, and we received 40 Post-It® notes with comments and suggestions. (See the Appendix for the draft Research Agenda and an image of the poster with notes.)

A research agenda can be used to integrate and support a sustained, systematic, and programmatic approach to conducting research about and for the archival profession. We believe that the Research Agenda is a living document that should be maintained by an active standing committee with reflection and input-gathering from SAA members interested in conducting, supporting, or otherwise making use of archival research.

The comments we received on the Research Agenda from 2018 Annual Meeting attendees provide important feedback on each of the five major categories. Below we summarize comments beside the category most relevant to the attendee’s concern.

- Diversity and Demographics: Institutional types. Mid-career progressions & entries (consider ageism). How many lone arrangers work without professional staff? Supporting small organizations (shifting statistics) per CLIR Hidden Collections; include government archives explicitly.
- Metrics and Institutions: Best teaching practices (for lone arrangers). “I’d like the ability / standard / benchmark to assess impact in a way that is also used by peers.” Impact metrics. Look at imbalances in approaching different domains of the archival profession. Clarify from measurement to action; what do our current descriptive standards miss and

how does that affect our analysis? Redo A*CENSUS, “sustainable” data about archives and archivists. Consider Data Rescue inventories and grants.

- Maintenance, Sustainability, and Ethics: The committee’s model (volunteer or paid staff? Source of funding). Data repositories for past surveys. Climate change.
- Inclusive Collaboration: Job descriptions and skill sets needed. Study archival users of traditional and digital data. Core competencies of accrediting agencies, distinguishing the archival uniqueness. Address the views of museum & library workers (and stop separating); use shared language and acknowledge shared goals. Build interdisciplinary work relating to the diversity, in which we are the experts. Comparative analysis of US, Canadian, UK, African, Slavic, Latinx archival practice (de-silo US theory). Computational archival science. Opportunities for public participation and engagement.
- Audience Building: Start archival high school & vocational technical programs. Advocacy. Record creators’ donations. For colleges & universities: think beyond the humanities. Publication trends, textual analysis of society’s public communications.

Following completion of the Research Agenda, the Task Force turned its attention to carrying out a pilot project to demonstrate the feasibility of a future body such as ours to conduct an organized program of research of a set duration. Below we describe our effort in that regard with our focus on the “Repository Data” project led by Eira Tansey and Ben Goldman.

Pilot Project: Repository Data

TF-CORDE was charged to actively pursue “one or two pilot projects as a means of 1) testing initial assumptions and 2) incrementally developing a potential operational model for a standing body.” We identified Eira Tansey’s and Ben Goldman’s Repository Data project as a pilot to explore two key areas:

- 1) How could an SAA committee support the deposit of and access to research data with significant long-term value to archivists for future research and the potential for re-use?
- 2) What levels of effort, expertise, and other infrastructural resources would be necessary for an SAA committee or SAA staff member to continue to maintain data about archival repositories around the U.S.?

Tansey and Goldman’s Repository Data project aims to comprehensively list all of the archival repositories in the United States ([as defined](#)) and to “map the vulnerabilities of American archives locations to the future impacts of climate change,” as a way to support responsible facilities management and rapid responses to potential crises (see [RBMS 2018](#)).

TF-CORDE saw this project as having fundamental significance to the future success of SAA research efforts, given the widespread interest in better comprehending and articulating the boundaries and makeup of our profession. We would like to first note the existence of two editions of *Directory of Archives and Manuscript Repositories in the United States* that were produced by the NHPRC in [1978](#) and [1988](#), which serve as important previous work for the research activities we explore below. Research questions that address the number of archivists, location of archives and type of archives provide core data with enduring value to SAA and the

archival profession. We suggest that such questions should continue to motivate the work of the future SAA committee.

In the pilot investigation, TF-CORDE examined the resource needs for stewarding the repository data and considered how its long-term stewardship would support and reinforce the research interests of SAA and its members. The Task Force focused first on how the Repository Data project team might generate a map of archives located in a U.S. region of interest: for instance, the [Gulf States](#). First, we understood how the process of gathering repositories for one state (Florida) revealed that data drawn from ArchiveGrid (currently a key finding aid aggregator for American archives) does need to be supplemented by contributions from knowledgeable or expert sources for each state; that is, where ArchiveGrid alone had indicated 22 repositories in Florida, the Repository Data team's [data gathering](#) over the last quarter of 2017 produced a count of 240 repositories, including many small and community archives not contributing finding aids here. (Similarly dramatic results for Ohio, from 44 to 500+ repositories, [are detailed here](#) as well as in [Archival Outlook](#), May/June 2018 issue.) Taken more broadly, the Repository Data team reported a count of 16,326 repositories in 31 states [in their April presentation](#). The important role of SAA's Regional Archival Associations Consortium ([RAAC](#)) and willingness of its leaders and members to participate in the work by Repository Data to expand our known universe of archival repositories for the regions they cover should continue to be duly noted.

The skills required to gather and compile the data provided by local archivists and RAAC members include: metadata normalization (selection and use of controlled vocabularies), categorizing institutions, deciphering physical from mailing addresses for purposes of visualizing map location, and publishing offline spreadsheet data for public access [in GitHub](#). The project initially sought to gather roughly 20 data fields for each archival repository in the U.S. (as detailed in the ["Data Decisions" blog post](#)) while maintaining fluency with archival standards such as DACS 5.6, RiC section 3, EAC, and Geocodio. Such work was performed by the Repository Data's grant-funded part-time (10 hours/week) research assistant over one academic year from late-September 2017 to late-April 2018, and is set to be completed for all 50 states by December 2018 by the two principal investigators.

The archival skills necessary to sustain and expand upon the Repository Data project can also be illustrated by a second exploration we undertook in anticipation of the potential path of Hurricane Florence (or "cone of uncertainty" encompassing four states) in mid-September 2018. From GitHub, the project's principal investigator input four states' repository data (see Figure 2) into ArcGIS online, adding three layers of information from the U.S. NOAA information service. The work to generate the resulting ArcGIS map is more fully detailed in a project blog post at: <https://repositorydata.wordpress.com/2018/09/12/tracking-hurricane-florence/>. Of note, the map provides a compelling visualization of the cultural heritage implications of the storm, and reinforces the opportunities Repository Data presents to us to raise awareness of the archival impact of climate change. The map can aid and elevate Coordinated Statewide Emergency Preparedness ([COSTEP](#)) leaders' communication with affected repositories.

Repository Name...	Name Notes	Parent Org	Authorized Repo...	Repository Ident...	Repository Type	Location Type	Street Address
Cave Spring Historical Society					Historical Society/Museum	Unverified	13 Cedartown
Cedartown Public Library	Branch of the Sara Hightower Regional Library				Public Library	Unverified	245 East Aver
Polk County Historical Society					Historical Society/Museum	Mailing Address	P.O. Box 203

Figure 2. Source data (some metadata fields) for the Repo Data map (13 Sept 2018) of archives in the potential path of Hurricane Florence – Georgia, South Carolina, North Carolina and Virginia. Map available at <https://arcg.is/uOizr>.

Outcomes of our pilot study of the Repository Data project include:

- Knowledge of the prospective process of SAA working with researchers to facilitate preservation and access to research data (gained primarily from: joining their conference calls, closely reading project documentation including blog posts, final report, and presentation materials). Process knowledge allows us to communicate particular research discoveries from the data (e.g. about the path of Hurricane Florence on archives) in a fuller context.
- Articulation of specific work time (10 hours/week over two semesters, Fall and Spring; specifically 300 hours from January-April 2018 on refining data received from 150 contributors), skill sets (responsive communication with RAAC representatives and project supervisors and interested audiences; data wrangling functions such as organization, normalization, and categorization; professional outreach and presentation including negotiating privacy issues), and tools (use of Excel, GitHub, ArcGIS, and Federal datasets) necessary for successful completion of the project and interim project outputs (e.g. storm maps).
- Descriptive and compelling details about how the Repository Data effort reinforces the long-term research interests of SAA and its members. Awareness of continuing professional concerns, such as over- or under-representation of institutional types in particular ways (e.g. standards-making and in ArchiveGrid) and persuasive visualizations, and moving from risk-quantifying to solution-framing.

TF-CORDE initially planned to work with the *American Archivist* editor Cal Lee on ingesting the data into the affiliated Odum Institute Data Archive; however, since the project is currently

ongoing, the data is not yet ready to be ingested into a new repository for sharing and re-use. We encourage both the project leaders and the future standing committee to identify the best solution(s) for ensuring the long-term availability of the data. The Odum Institute's [workflow information](#) provides valuable guidance to researchers about ensuring future data availability.

In summary, we articulated two sets of needs that emerged from our pilot project work (one goal and one process):

1. To preserve data for anyone to use, and
2. To figure out, with specifics, how SAA can take direct responsibility for the data and continue to maintain accurate information about archival repositories. The committee can investigate possibilities for collaborating with other entities for maintenance of repository data.

Based on our pilot investigation here, we suggest that the future SAA committee endeavor to take a similarly detailed, case-by-case approach with this or another similar research project(s) in order to support its far-reaching research and information potential. As a task force, for example, we agreed that one reason to investigate Repository Data as our pilot is because its research products (maps and visualizations) are immediately useful, the overall project is nimble and responsive to current events, and it demonstrates how SAA and archival research can be consequential and impactful. Repository Data also helped our task force understand the requirements, responsibilities, and activities of a committee seeking to ensure the preservation of such data (whether in GitHub or in combination with a hosted repository) and enable its expansion, access, and use long-term. In considering the future of the Repository Data project, we were reminded of an analogy case—the [SAA Glossary](#), which started as the personal project of SAA member Richard Pearce-Moses before becoming the province of SAA and, currently, the SAA Dictionary Working Group. We would like the future committee to consider similar arrangements for collaborating with the project leaders to maintain archival repository data.

Based on our work with the Repository Data effort, we offer two future directions worth pursuing:

1. To integrate in some form with the National Inventory of Humanities Organizations ([NIHO](#)), which is a similarly new project of the American Academy of Arts & Sciences. We note that while Repository Data currently collects information about “repository type” (e.g. historical society, public library, K-12, government), the project currently does not collect uniform information on repositories’ MARC repository code (if one exists) or Employer Identification Number (EIN), for its repositories. We suggest that the future SAA committee might consider ways to supplement the existing project data with these additional data points in order to pursue broader goals. In addition to NIHO, there are related datasets maintained by IMLS ([Museum Universe Data File](#)), [NEH](#), and [NHPRC \(1978 and 1988 Directories\)](#) about libraries, archives, museums, and galleries (GLAMs) that further expand its potential.
2. Another future direction is to further explore data deposit with the Odum Institute, in collaboration with the *American Archivist* Editor Cal Lee. We encourage a formal articulation of the purpose and benefits of depositing a dataset into the “dataverse of preference” of the *American Archivist* (e.g. is table form best?). As we mentioned above,

we considered a practice ingest of project data during our time as a task force, but timing and the need to study a sample deposit agreement (the implications of and ability to update an ingest) both suggest that such work would be better handled as a responsibility of the future committee. We did not wish to make assumptions about data ownership before they are appropriate and before the committee has laid out the issues (e.g. Creative Commons and copyright) fully.

Overall, we recommend that the future committee center the archival value of transparency in deciding and determining future research and data activities. Here we mean transparency with regard to SAA's approach and principles for data decision-making. Part of such transparency in practice will acknowledge data reuse and data ownership as issues that should be agreed to in writing, acknowledging myriad possible uses of data by many users. The committee should consider SAA's relationship to data it may not own but wants to provide access to, and be transparent about the stance taken regarding protection, preservation, and delivery of appropriate items, but also ensure it is freely accessible (for example, ARL data is behind paywall, ARMA offers a salary survey report for \$90) beyond members-only as determined. The SAA Council should acknowledge and discuss various business model options regarding access to data, as a possible SAA member benefit as well as the broader goals of sharing archives research data (absorbing and recouping costs of producing the data). Finally, the future committee should also look for opportunities to broadly communicate particular baseline data about archives and archivists.

CONCLUSION

TF-CORDE recommends the establishment of a Committee on Research, Data, and Assessment (CoRDA), and has identified four primary focus areas for this standing committee:

- 1) Collect and analyze data in support of SAA's key strategic goals;
- 2) Help archivists find and share data to support their work;
- 3) Help archivists learn how to collect, manage, and use data effectively for their work; and
- 4) Identify and pursue funding and/or collaborative opportunities to increase the gathering, publication, and use of data about archives and archivists.

The task force recommends establishing the committee as a scalable entity that is initially driven by volunteer efforts and lightly supported by existing staff. Similar cultural heritage professional organizations that have dedicated staff, often 1+ FTE, also have a much larger membership base than SAA and therefore potentially more funding for such activity. Dedicating staff to this effort is just not feasible currently.

The initial goals of the Committee will be to:

- Investigate the means of conducting an annual archival demographics survey and subsequently making its results available;
- Further explore data repository (deposit) options and version updates;
- Work with SAA staff to enhance/convert the [“Facts & Figures” page](#) into more of a research, data, and assessment information portal;

- Collaborate with the SAA Education Department and the Committee on Education to identify data collection and analysis gaps and educational opportunities; and
- Investigate funding models for a more robust research, data, and assessment agenda.

RECOMMENDATIONS

THAT a standing committee on Research, Data, and Assessment be established, per the following description:

**Committee on Research, Data, and Assessment
Description and Charge**

I. Purpose

The Committee on Research, Data, and Assessment provides access to significant and useful data and research about SAA, American archives, and their users, which evidence the value of archives for society and help us improve our services to SAA members and to our consumers. The committee will work to conduct or support relevant research and to create, gather, and preserve data by directing and engaging in several areas of activity:

- Providing SAA members with standardized tools for gathering and analyzing data;
- Providing a repository or portal for data and other research outputs;
- Providing training on gathering, analyzing, interpreting, and using data; and
- Providing up-to-date and reliable basic facts and figures about archives and archivists.

II. Committee Selection, Size, and Length of Term

The committee shall consist of a minimum of nine appointed members (including two co-chairs) serving staggered three-year terms with the possibility of reappointment. The Vice President, on behalf of the Council and with the recommendation of the committee, appoints new members and co-chairs. Since committee members may need to possess specialized skills and abilities, which will vary from time to time, an RFA (Request for Applicant)-based process may be necessary to help surface the best candidates in any particular appointment cycle. Maximum committee size is variable, dependent upon the number, nature, and complexity of the projects and activities in which the committee is engaged at any particular time.

To better facilitate the committee’s diverse work, members may be distributed among multiple project- or activity-based subcommittees, each of which is headed by a subcommittee chair appointed by the committee co-chairs to repeatable annual terms. The Committee co-chairs may recommend that the SAA Council form (and disband) subcommittees from time to time, based on current needs.

The vice chair of the Committee on Education will serve as an *ex officio* member of the committee.

III. Reporting Procedures

The committee co-chairs shall submit a report for each Council meeting, summarizing current activities and projects and progress toward stated objectives. As directed by the Council, the committee chair may submit a written report upon the conclusion of specific projects.

IV. Duties and Responsibilities

The committee shall enjoy a great deal of latitude in developing and engaging in research projects and data accumulation. In so doing, the committee must maintain frequent communication with the Council, which approves and advises on the committee's strategic directions, tactics, resources, activities, and projects. This communication is facilitated by the committee co-chairs, who must ensure that the Council is aware of emergent planning and directional changes.

The Committee maintains responsibility for achieving its purpose by::

- Looking strategically at the organization's information needs and sharing its strategies with the SAA Council and relevant component groups;
- Proactively commissioning or directly engaging in necessary and strategic research;
- Gathering quantitative and qualitative information of strategic value for SAA in the areas of advocacy, public awareness, improved audience service, and community engagement;
- Evaluating such information and surfacing it to SAA members via dashboards, reports, and constructed data sets that members can use to better understand and act upon their own environments and to perform better as archivists;
- Engaging academic and other communities within SAA to perform research by helping to set research agendas and by utilizing grants, fellowships, conferences, and other tools and levers;
- Providing a repository (or repositories) and analytical tools for sharing and evaluating useful data about archivists, repositories, audiences, and the environments in which archives function; and
- Building and maintaining, in collaboration with other SAA groups, a training site for archivists in the area of research, data gathering, evaluation and assessment, and business intelligence.

V. Meetings

The committee shall conduct its business largely through email and conference calls. The committee shall meet formally each year at the SAA Annual Meeting. Any additional in-person meetings that may be necessary shall only occur pending Council approval and funding.

Support Statement: The TF-CORDE, having investigated the questions posed in its charge to the best of its ability, has determined that a standing Committee on Research, Data, and Assessment would be useful to SAA and its members and would advance important elements of the SAA strategic plan. If the committee is constituted in a scalable manner that does not put a new and immediate burden on the association's resources, it is possible to establish the committee in the near future.

Impact on Strategic Priorities: Establishing a standing committee would support and help advance three different priorities within the 2018-20 SAA Strategic Plan:

- 1.4. Strengthen the ability of those who manage and use archival materials to articulate the value of archives.
- 3.1. Identify the need for new standards, guidelines, and best practices and lead or participate in their development.
- 3.2. Foster and disseminate research in and about the field.

Fiscal Impact: A standing committee has the potential to consume significant operating resources, including the allocation of new staff. Therefore, the task force recommends establishing the committee as a scalable entity that is initially driven by volunteer efforts, lightly supported by existing staff, and with the possibility of some projects supported by the SAA Foundation and external grants and gifts. Additional resources may be allocated as the committee demonstrates the value of its efforts.

Appendix

A Research Agenda for Leading the Archival Profession

Archivists clamor for a sustained, systematic, and programmatic approach to conducting research about and for the archival profession - as seen by the multitude of independent and SAA Section-based research activities launched every season and the subsequent appearance of those research results and supporting data in just as many venues and outlets. There is wide acknowledgement that the A*CENSUS, funded by the U.S. IMLS and fielded in 2004, was a landmark research study that obtained important and original insights about the nature of archival work, archival education (graduate and continuing), and demographics. We are interested in supporting the next iteration of profession-wide research by outlining major topics and themes that can attract teams to pursue projects now and into the future, nationwide. Our Task Force offers the following research agenda in order to capture the areas of interests and necessary directions expressed both in the course of interviews we conducted (for the purpose of persona development) in 2017-2018 and from key literature sources concerning archival research. A future SAA Committee on Research, Data, and Assessment (CoRDA) should endeavor to conduct research around:

Diversity and Demographics: Support broad participation in archival research activities

- clearinghouse for launching and learning from SAA component group-driven surveys and projects (retrospective and continuing)
- Assessment and evaluation of graduate and continuing archival education, on such topics as syllabi analysis and curriculum development, cohort demographics (e.g. 1998 KALIPER Project, AERI Initiative), recruitment and retention in the profession, professional development resources (SAA Education+ offerings)
- Statistics about the archival profession and longitudinal data-collecting
- Term employment and its effects; skilled labor in the profession
- Accessibility in the archives
- [PNAAM](#), Indigenous and tribal archives
- Archival leadership and managing career paths and team transitions
- Archival advocacy and its teaching and practice; born-digital gap analysis

Metrics and Institutions: Support efforts to articulate explicit value propositions for special collections and archives

- collection metrics (anticipated 2018, see [SAA/RBMS Task Force](#))
- usage metrics (see SAA/RBMS Task Force's [Standard, 2018](#))
- State Archives and data reporting infrastructures
- Universe of archival repositories (e.g. NHPRC *Directory* [1978](#), [1988](#); OCLC ArchiveGrid, [Repo Data](#), [Lavender Legacies Guide](#))
- From descriptive data to actionable goals for data (e.g. [ARL Statistics](#))
- Performance measures and service benchmarks (analyzing current practices geared toward future recommendations)
- Appraisal and the ecosystem of collection formats and data (manuscripts, web, audio-visual, artifacts, post-custodial collecting)

Maintenance, Sustainability, and Ethics: Continue to conduct needs assessments sustainably

- Strengthening collections infrastructure (e.g. preservation documentation, Project ARCC)
- justifying particular archival activities: While salary surveys have an important role, some members are also interested in understanding particular components of archivists' work, like practices (description, VIP tours, management scope, [recruitment criteria](#)), collection surveys, and facilities. Often to make a stronger case to upper administration but also for professional development and educating/training others. e.g. [Senate Archivists](#), Congressional Papers (digital) archivists ([2009](#)), film archivists ([2015](#)), museum archivists ([2010](#), [2014](#)), new archivists ([1](#), [2](#), [3](#)), Public Library archivists ([2013](#)), accessioning archivists ([2011](#)), and [would-be /curious](#) archivists.
- User-centered and data-informed research studies to support public investment in archives' public services and stewardship
- Standards-making support and maintenance (discovery systems, rights toolkit, [collections as data](#), information governance)
- Open-source and open-access collections (management and development) (e.g. Greg Eow's 2017 plenary, EAD aggregators, academic research datasets)

Inclusive Collaboration: Act in the spirit of the SAA GPAS curricular component of ["complementary knowledge"](#) to expand research with allied organizations

- Inclusion in K-12 curriculum and teacher training with primary sources
- School library partnerships for school archives and commemorative events
- Spotighting archives in Home Movie Days and National History Day competitions
- Community-based archives, family archives and genealogical research, military archives - collaborative digitization programs
- Community service and knowledge contributions of SAA Student Chapters (e.g. mini-conferences, campus initiatives)
- Competency-oriented archival education and training (e.g. [ACRL Competencies for Special Collections Professionals](#))
- Acknowledging common purpose with galleries, libraries, and museums (GLAM) and deep problem-solving with regard to prospective digital shifts, environmental footprints and climate change, and subject and language expertise (e.g. the work of [CALM](#) and [Collective Wisdom](#) exchange, [2016 ARL-CNI-Miami Summit](#))
- Activism and imaginaries: human rights archives, anonymity and data privacy
- Archives as/for reckoning (e.g. Archival Platform, in higher education contexts: Project STAND, Slavery and Justice)
- Communication across repository types (e.g. [RAAC](#))

Audience Building: Broadening archival awareness, funding, and partnerships

- regularly update the SAA [Facts & Figures](#) page (e.g. with Section surveys)
- Nurture relationships with media outlets (e.g. CoPA ArchivesAWARE)
- Promote the global profession (e.g. #AskAnArchivist)
- Biographical research and leadership storytelling (e.g. 75th Anniversary trading cards and oral histories, Harold T. Pinkett's legacy, Mosaic Scholars, [STHC "Unsung Heroes"](#))
- The history of archival ideas and consciousness (e.g. ancient archives)

- Managing up: (campus) case studies of success and lessons learned (e.g. [Jump In!](#) initiative)
- Multilingual corporate marketing and sources of revenue-generation
- Cultural heritage and virtual heritage initiatives (e.g. SAA [CHWG](#))
- Cross-disciplinary grand challenges: peace building, green facilities / green economy, immigration, accountability and transparency, ways of recordkeeping, research and/as practice

References

Archival Platform (2018). University of Cape Town. <http://www.archivalplatform.org/about/>

Arnold, H. (2017, Mar 24). Standards work: Revising the DACS Principles (in a blizzard). *Bits & Bytes: News from Rockefeller Archive Center's Digital Team*. <http://blog.rockarch.org/?p=1710>

Cifor, M., Evans, J., Punzalan, R., Ramirez, M.H., Soyka, H. Wartenbe, M., & Wilczek, E. (2014, July 14). "Grand challenges." Plenary panel at AERI 2014, University of Pittsburgh. http://www.ischool.pitt.edu/aeri2014/documents/grandchallenge_aeri2014_2014-07-14.pdf

Coelho, C. (2018). Anniversary in Review 2014-2015. *Brown University, Center for the Study of Slavery and Justice*. <http://250.brown.edu/anniversary-in-review>

Council of State Archivists. (2013). *The importance of state archives*. Albany, NY: CoSA. https://www.statearchivists.org/download_file/view/321/2123/

Cox, R.J., Yakel, E., Wallace, D., Bastian, J.A., & Marshall, J. (2001). Archival education in North American library and information science schools. *Library Quarterly: Information, Community, Policy* 71(2):141-194. <https://doi.org/10.1086/603260>

Dooley, J. M., & Luce, K. (2010). Taking our pulse: The OCLC Research survey of special collections and archives. Dublin, OH: OCLC Research. <http://www.oclc.org/research/themes/research-collections/hiddencollections.html>

Duff, W., Yakel, E., Tibbo H., Cherry, J., Krause, M., & Sheffield, R. (2010). The development, testing, and evaluation of the Archival Metrics Toolkits. *American Archivist*, 73(2), 569-599. <https://doi.org/10.17723/aarc.73.2.00101k28200838k4>

Dupont, C. and Yakel, E. (2013). "What's so special about Special Collections?" Or, assessing the value special collections bring to academic libraries. *Evidence Based Library and Information Practice* 8(2): <https://journals.library.ualberta.ca/ebliip/index.php/EBLIP/article/view/19615/15221>

Mazurczyk, T., Piekielek, N., Tansey, E., & Goldman, B. (2018). American archives and climate change: Risks and adaptation. *Climate Risk Management*: <https://doi.org/10.1016/j.crm.2018.03.005> and see also <https://repositorydata.wordpress.com/>

McKemmish, S., with Gilliland, A., Evans, J., Lau, A., & Rolan, G. (2016, July 9). "ARK research: The state of the art." Plenary presented at AERI 2016, Kent State University.

Meissner, D. (2014, Sept 25). Archives, archivists and data gathering. *Off the Record*. <https://offtherecord.archivists.org/2014/09/25/archives-archivists-and-data-gathering/>

Meissner, D. (2017, Spring/Summer). Bare necessities. *American Archivist* 80(1): 6-18. <https://doi.org/10.17723/0360-9081.80.1.6>

Project_ARCC: Archivists Responding to Climate Change. (2018). <https://projectarcc.org/project-marcc/>

Project STAND: Student Activism Now Documented. (2018). <https://standarchives.wordpress.com/>

SAA Section Feedback. (2017). Excel file provided to TF-CORDE by SAA Staff.

SAA Strategic Plan. (2017). <https://www2.archivists.org/governance/strategic-plan>

Weber, C. S. 2017. *Research and learning agenda for archives, special, and distinctive collections in research libraries*. Dublin, OH: OCLC Research. <http://doi.org/10.25333/C3C34F>

Zanish-Belcher, T. (2018, April 23). Some archives questions need answers. *Off the Record*. <https://offtherecord.archivists.org/2018/04/23/some-archives-questions-need-answers/>

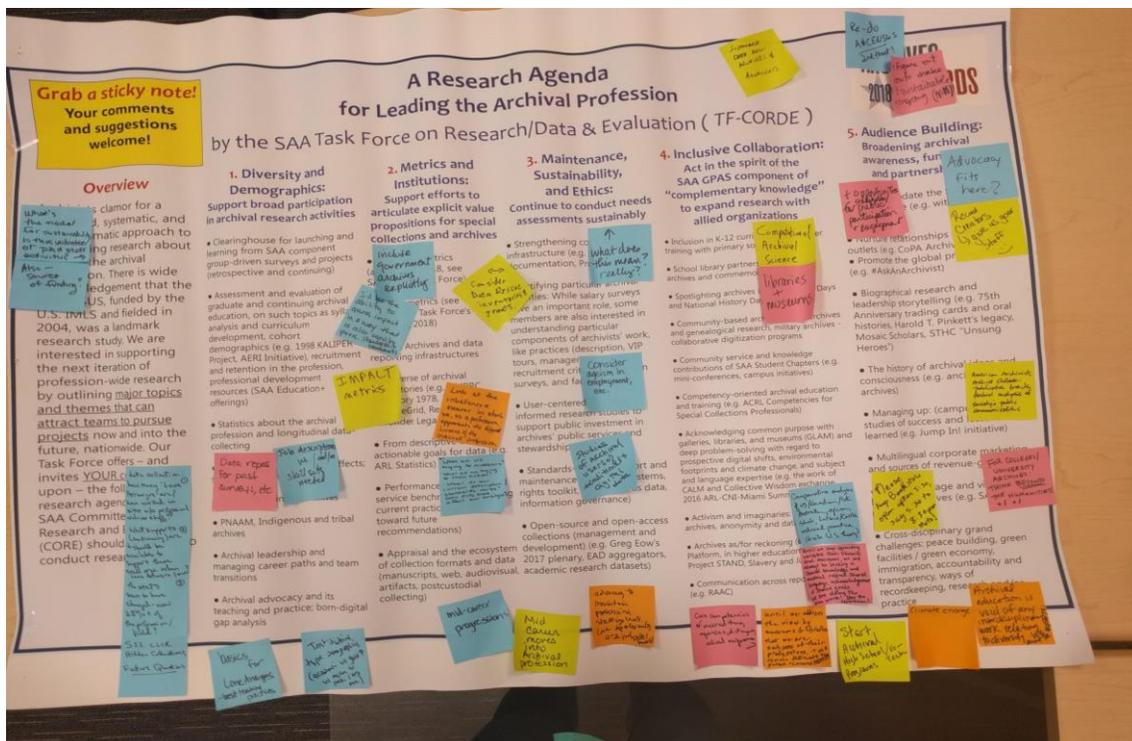


Figure 1. SAA 2018 attendee comments on our Research Agenda.

Available at <https://drive.google.com/open?id=1kfZ9Bx2IBttvrMeqVD55V6pRBxC9KkTL>