**SAA Committee on Research, Data, and Assessment (CORDA)**

**SAA Dataverse Collection Development Policy**

**February 24, 2021**

**Statement of Purpose**

The Society of American Archivists (SAA) Dataverse is an SAA data service that was established to support the needs and interests of SAA's members and the broader archives community. The SAA Dataverse supports the reuse of datasets for purposes of fostering knowledge, insights, and a deeper understanding of archival organizations, the status of archivists, and the impact of archives and archival work on the broader society. Deposited datasets should be “actionable” in that they should support direct analysis and interpretation.

The SAA Dataverse welcomes deposits of collections of quantitative or qualitative data and associated documentation.

**Guiding Principles and Collecting Focus**

* The SAA Dataverse supports active data reuse and reevaluation on behalf of the archival community.
* The SAA Dataverse will adhere to [FAIR Guiding Principles for Scientific Data Management and Stewardship](https://www.go-fair.org/fair-principles/) (Findable, Accessible, Interoperable, and Reusable).
* The SAA Dataverse systems, policies, and procedures align with standards for trustworthy digital repositories, providing long-term data archiving, bit-level preservation, and data sharing that foster analysis, reuse, and new interpretations.
* The scope of the data and associated documentation must have clear potential value to one or more of four designated communities:
	1. Archives and records professions in North America
	2. Society of American Archivists
	3. Archival repositories
	4. Individual professional archivists and archival scholars

The SAA Dataverse welcomes data created by the SAA community and allied organizations and individuals.

* SAA requires that SAA-sponsored data be submitted for deposit. This includes data gathered by the Council, component groups, work supported by SAA Foundation funding, and selected materials related to the SAA Research Forum
* Archival repositories that gather data of interest and usefulness beyond the originating organization
* Individual professional archivists and archival scholars who gather data of broad interest and value
* Authors publishing research in the American Archivist, a peer-reviewed journal published by the Society of American Archivists.

SAA will prioritize data identified by and solicited for deposit by CORDA in its role as developer and manager of the SAA Dataverse repository.

**Submission Requirements and Data Formats:**

Datasets and associated documentation must conform to the SAA Dataverse Terms of Use and SAA Dataverse Data Deposit Form. (link once ready). Submissions will be reviewed by a designated review committee of SAA’s Committee on Research, Data, and Analysis (CORDA), and ingest will be depend on adherence to the Guiding Principles and Collecting Focus. Review does not include assessment of methodology used.

Data should provide information about:

* Origins of the dataset;
* Methodologies for gathering, preparing, and analyzing the data;
* Any analyses or reports that are derived from the data, either published or unpublished.

Associated documentation can take the form of:

* Tabular data (including Excel, CSV, R, Stata, SPSS, tab-delimited files)
* Image files (.jpg, .png. .tif)
* Geospatial files
* File types that are widely accepted, that can be converted, and that are non-proprietary/open source (PDF, etc.).

All contextual information necessary to understand and use the data (such as readme files, checklists, etc.) is a mandatory component of the dataset. The submitter must remove all personally identifiable information (PID). The submitter must supply metadata to the fullest extent possible at the time of ingestion.

**Preservation:**

The SAA Dataverse is hosted by the Odum Institute for Research in Social Science at the University of North Carolina at Chapel Hill, which is a trusted repository for research data, datasets, and associated research whose mission is to provide long-term data archiving and data sharing. The Odum Institute’s systems, policies, and procedures have been developed in alignment with standards for trustworthy digital repositories as outlined in [ISO 14721 Reference Model for an Open Archival Information System (OAIS)](http://www.oais.info/) and ISO 16363 Audit and Certification of Trustworthy Digital Repositories. Data is preserved at the bit-level. Deposited items will be retained for the lifetime of the SAA Dataverse.

**Copyright and Intellectual Property Rights:**

All datasets and supporting documentation must be granted for distribution in the SAA Dataverse. Uploaded materials must be made freely available via a Creative Commons license. Datasets and associated documentation may be deposited in repositories outside of the SAA Dataverse if permission and copyright allow.

**Deaccessioning:**

Deaccessioning records from the SAA Dataverse should occur only if there is a legal or otherwise valid reason for the dataset to no longer be accessible to the public. If a record must be deaccessioned, a version of the dataset or the entire dataset may be deaccessioned. If a record was published for any amount of time, a “tombstone page” will remain for records or data that have been deaccessioned to indicate its removal. Submitters must work with CORDA to deaccession data or records.