

The American Congress Digital Archives Portal Project

December 2023

Prepared by

Danielle Emerling, <https://orcid.org/0000-0002-8525-491X>

Sarah Gard, <https://orcid.org/0009-0002-2691-0638>

Jessica McMillen, <https://orcid.org/0009-0001-9049-6649>

Marla Weaver, <https://orcid.org/0000-0002-9020-4781>

This project was made possible in part by a 2022 award from the Catalyst Fund at LYRASIS. This work is licensed under Creative Commons Attribution-NonCommercial-ShareALike 4.0 International License.

PROJECT BACKGROUND

The archives of the United States Congress document the democratic process, the development of public policy, and the evolution of the legislative branch. Held at institutions throughout the country, these materials constitute part of the history of the largest and most diverse, but arguably least understood, branch of U.S. government. Congressional archives are vitally important for understanding when and how the body wields its powers, which reflect and shape the character of America. Additionally, the value of these archives goes beyond the study of the branch itself, as they illustrate multiple narratives related to the country's social, cultural, and political development. The American Congress Digital Archives Portal (ACDAP) project aims to provide open access to congressional archives by bringing together these civically important sources from multiple institutions using open-source software. Congressional materials in a single online portal will illuminate the value of each collection and the relationships among them, as well as provide opportunities for new scholarship and civics and history education.

Congressional records are created by numerous collaborating entities, such as the executive and judiciary branches, constituents, and outside interest groups. The primary materials generated by Congress are separated into official and private records. The official records of Congress are those created, received, and maintained by committees, and they remain in the custody of the National Archives and Records Administration. The official records are differentiated from the personal papers of Members of Congress. Personal papers include materials created or received by the individual Member as documentation of his or her career, the organization and functions of the office, and the work of office staff.¹

Because Members of Congress own their personal papers, Members decide where to place their collections. Members have donated papers to large and small academic institutions, historical societies, state archives, and to named centers, which usually focus on the study of policy and politics. It is the personal papers of Members of Congress that this project seeks to make available. Personal congressional collections are most often named for the Member of Congress, but this obscures the true breadth of content and the many formats they contain. Collection contents include materials related to legislation, nominations, and investigations; press relations; personal and political activities; and constituent services. Legislative files reveal decision making and the often hidden and complex policy making process on issues both foreign and domestic. Constituent services records contain the voices of people who are rarely documented in the archives, expressing their views on issues important to them and seeking assistance with government programs. Political and press files are instructive for understanding parties, influence, and the ways in which media shape discourse. Materials are national,

¹ Karen Dawley Paul, *The Documentation of Congress: Report of the Congressional Archivists Roundtable Task Force on Congressional Documentation* (Washington, DC: Government Printing Office, 1992), 102-20, and "Congressional Papers and Committee Records: Private vs. Public Ownership," in *An American Political Archives Reader*, eds. Karen Dawley Paul, Glenn R. Gray, and Rebecca Johnson Melvin (Lanham, MD: Scarecrow, 2009), 91-94.

regional, and local in scope. From sweeping national policy to local public projects, they document how, what, and why our society allocates its resources and expresses its values.

Congressional archives provide an essential underpinning for scholarship and have been used to examine, among other things, race, gender, and ethnicity in relation to domestic programs; the expansion of federal government; American conservatism; economics and capitalism; history of the institution and its members; and Congress's role in larger narratives of different historical periods.² They have been essential for the study of congressional procedures and leadership, relationships with the other branches, and how public policy shapes national life and social relationships.

However, scholarship about Congress and its role in shaping American life and policy has been less robust than would be expected. Further, the stories and contributions of underrepresented groups—women, Black Americans, Hispanic Americans, Indigenous Americans, and others—to congressional history and public policy remain even more obscured. In part this gap is due to the complexity of Congress, the large number of individuals and committees involved in policy making and shifts in scholarly interests and trends. Practical barriers to using congressional archives also exist. Unlike presidential papers, which are centralized in one location with dedicated staff and funding, congressional collections are geographically dispersed among institutions large and small with varying degrees of resources for processing and providing access to the collections. The collections themselves are often extremely large and complex, demanding several years of archival processing. For scholars, collections may be difficult to use, both because of a lack of travel funding and the breadth and varying levels of description in congressional archives. The COVID-19 pandemic exacerbated the issues related to using archives due to travel restrictions, archives reading room closures, and limited or non-existent research funds.

The ACDAP project remedies these problems by digitizing selected congressional archives and aggregating them into a single portal, providing scholars with significantly improved and more equitable access to geographically dispersed collections and providing the public and educators with access to civically important documents about the legislative branch and American public policy.

A pilot (Phase 1) of the project received a Humanities Collections and Reference Resources foundations grant from the National Endowment for the Humanities (NEH), 2021-2022. During this pilot project, West Virginia University Libraries (WVUL) partnered with the Robert C. Byrd Center for Congressional History and Education and the Robert J. Dole Institute for Politics at the University of Kansas to develop a portal prototype, perform a user survey to guide content selection, and conduct usability testing to improve portal functionality. It also established a metadata model and best practices related to managing copyright, privacy, and classified materials. The user survey and user testing revealed that potential users saw great

² Julian Zelizer, "Congress is Back: Scholars Study its History to Understand its Problems," *The Chronicle of Higher Education*, June 2013.

promise in the project and desired easier access to congressional archives, as well as their preferred methods for searching, faceting, and exploring collection materials.

Phase 1 also included planning for governance and growth. It established a Memorandum of Agreement among partners and included initial work toward long-term management structures; the portal groundwork was laid with one release with future plans for expansion; funding was secured from the NEH and the Association of Centers for the Study of Congress (ACSC); and engagement with the congressional archives community was frequent through organizations like the ACSC and the Society of American Archivists' Congressional Papers Section. A [white paper](#) documented the work of Phase 1 in detail.

PROJECT OBJECTIVES AND TIMELINE

With support from the Catalyst Fund, the project moved into Phase 2. WVUL partnered with the Robert J. Dole Institute of Politics at the University of Kansas and development company SoftServ to carry out three overarching objectives from July 2022-June 2023.

Objective 1: Improve and reconfigure the underlying technical infrastructure.

Partners determined that to scale the project, the application would need to transition from a content hosting to a harvesting model. In Phase 1, WVUL hosted partner content using a Samvera/Fedora stack. In Phase 2, WVUL developers and web application development company SoftServ (formerly Notch8) planned to build on that stack, contributing any new development knowledge back to the community. They would also deploy the portal application from WVUL internal servers to Amazon Web Services (AWS) so that the project could begin to stand on its own without significant support from a host institution. Additionally, the metadata model would be adjusted to allow for harvesting digital object metadata and thumbnails from partner institutions.

Objective 2: Create a roadmap for the project's Diversity, Equity, and Inclusion initiative.

In the 2022 project white paper, partners laid out plans for the project that included focusing on the contributions underrepresented people and communities have made in Congress by including archives of women, Hispanic Americans, Black Americans, Asian and Pacific Islander Americans, Indigenous Americans, LGBTQ+ Americans, and Americans with disabilities who have served in Congress. The project would use the help of a consultant to create a set of guidelines and actionable steps to act as a roadmap for engaging with potential partners, archival collections, and user communities.

Objective 3: Position the project for expansion.

Finally, the work in Phase 2 would help position the project to compete for funding, prepare to onboard additional partners, and ingest digital archives from those partners. Five

partner institutions had committed to apply with WVUL for an NEH implementation grant. These included the Dole Institute, the Dirksen Congressional Center (Illinois), the Carl Albert Center at the University of Oklahoma, the Richard B. Russell Library for Political Research and Studies at the University of Georgia, and the Hawai'i Congressional Papers Collection at the University of Hawai'i. These new partners would bring a diversity of collections and materials to the project and would help to fulfill its research and education mission.

PROJECT WORK AND OUTCOMES

In Phase 2, partners generally adhered to the timeline and achieved most of the project's objectives. A detailed explanation of outcomes for each objective follows.

Objective 1: Improve and reconfigure the underlying technical infrastructure.

In August and September 2022, WVUL investigated moving the application and continuing development on AWS, but initial and ongoing subscription costs were higher than those estimated when the proposal was written, and thus prohibitive. With the permission of WVUL administration, WVUL and SoftServ developers pivoted to continue hosting and development on WVUL servers. This required some additional administrative work to provide access and permissions to the SoftServ developers but was ultimately a workable solution. SoftServ developers configured harvesting for externally hosted content, requiring a base configuration of an importer. In this model, partner institutions host their own materials online, and the project administrators ingest metadata (via CSV) using the application.

WVUL worked with the Dole Institute to modify the existing metadata model to better meet the needs of the harvesting model. This process took a great deal of time and many conversations among partners and others. The project is grateful to Emily Fidelman, Head, Metadata Services at WVUL; Michelle Echols, Metadata and Cataloging Librarian at WVUL; and Jeremy Bartczak, Manager, Metadata Discovery Services at the University of Virginia Library for their time and thoughtful answers to questions about schema and aggregation. Major changes to the model included adding fields for previewing content (`edm:preview`) and viewing the digital object with metadata (`edm:isShownAt`). These fields were added to harvest thumbnails of objects and provide connections to the original object. A record type label (`dcterms:type`) was added so the application can render the material in the correct format. In response to the user surveys conducted in Phase 1, one field, Physical Location (`dcterms:isPartOf`), was added to create better connections between digital objects and the originating collection. Institutions will populate this field with collection, box, and folder information. To facilitate better date faceting, a second date field (`dcterms:created`) was added and will utilize the Extended Data/Time Format (EDTF) Specification. Another date field, which is human readable, is also a part of the model. (See Appendix B: Metadata Model)

The Dole Institute used the new model to create a test ingest case of approximately 20 digital objects in various formats, including text, photographs, audio, and video, and from a

variety of online sources where digital objects were made available. These included YouTube, Omeka.net, as well as those hosted on the Institute's finding aid server. The ingest was an iterative process as problems were addressed and metadata specifications were refined.

One significant challenge was rendering and displaying digital object thumbnails. In Phase 1, with the Portal hosting digital objects, the display of thumbnails was tied to the ingest process into the underlying system. In the harvesting model, the thumbnails for the objects were not displaying after a Bulkx import. As mentioned above, the field `dcterms:type` was added to enable the system to create a thumbnail image based on the type of item (e.g., photograph, document, video, audio, etc.). The WVUL systems development team updated the system to create thumbnails from the links provided in the metadata based on these item types. Because several of the objects were hosted in different repositories and systems, some customization was done to ensure thumbnails were created and displayed for all possible items. For items where it was not possible to create and display a thumbnail, the team created a placeholder image for each item type that gives visual continuity to the portal.

Objective 2: Create a roadmap for the project's Diversity, Equity, and Inclusion initiative.

Partners drafted and circulated a call for applications for a DEIA consultant in fall 2022. The consultant would draft a set of guidelines and actionable steps to act as a roadmap for engagement with potential partners, archival collections, and user communities. Partners may be institutions that preserve the collections of Members of Congress who identify as part of an underrepresented group. Archival collections may be created by a member of an underrepresented group or may contain significant materials related to an underrepresented group. Communities may be scholars, students, or others who traditionally do not use congressional archives but who may, through their use, contribute new perspectives and understanding about public policy, American history, or Congress as an institution. The partners engaged in a few promising conversations with a consultant, but unfortunately due to personal matters, the consultant was unable to complete work on the project. (See Appendix C: DEIA call for proposals).

Objective 3: Position the project for expansion.

The project partners reviewed the "It Takes a Village" Toolkit and identified steps that have been completed or need to be done in the areas of governance, technology, and community engagement. This provides a valuable framework for partners to work from. (See Appendix D: ITAV Toolkit Checklist). As Phase 2 came to a close, partners learned the project would receive an implementation award from the National Endowment for the Humanities. The NEH award will fund the project through June 2025. WVUL also learned in July 2023 that a congressionally directed spending (CDS) request for \$1.5 million for the project had been approved and included in an appropriations bill of the U.S. Senate Committee on Appropriations. While the CDS request awaits final passage, project partners are implementing the work plan for the NEH grant and planning next phases.

NEXT STEPS

As mentioned above, Phase 1 (2021-2022) of the project involved building a prototype and collecting information from stakeholders, and Phase 2 (2022-2023) focused on application development, DEIA work, and preparation for expansion.

Phase 3 (2023-2025)

In Phase 3, funded by an award from NEH (2023-2025), the project will add partner institutions, harvest thousands of materials from partners' collections, improve the Portal application, and improve the user experience. Partner institutions include the Carl Albert Congressional Research and Studies Center at the University of Oklahoma, Everett McKinley Dirksen Congressional Leadership Research Center, the Richard B. Russell Library for Political Research and Studies at the University of Georgia, the Robert J. Dole Institute of Politics at the University of Kansas, the University of Hawai'i Congressional Papers Collection, and West Virginia University. Partner institutions will digitize and/or aggregate materials totaling approximately 102,000 pages, 11,200 photographs, and 24,500 audiovisual items. Working with SoftServ and WVU developers, the project will improve search and facet customization, advanced search, metadata exports from searches, and citation generation. The partners will create research **guides to aid users in finding and understanding content.**

The project in this phase will also plan for future governance of the project and document onboarding processes for new partner institutions. Finally, the project will initiate some public-facing scholarship and civics and history education programs that use materials from the Portal.

Phase 4 (2025-forward)

In Phase 4, the project will continue to add partners and content and support scholarship and education initiatives. Development of the portal application will transition to maintenance. The most important part of this phase will be implementing a governance plan to sustain the Portal into the future.

APPENDICES

Appendix A: Project Team

Danielle Emerling (Project Director), West Virginia University

Sarah Gard, Robert J. Dole Institute of Politics, The University of Kansas

Steven Giessler, West Virginia University

Tracy McCormick, West Virginia University

Jessica McMillen, West Virginia University

Marla Weaver, Robert J. Dole Institute of Politics, The University of Kansas

Appendix B: Metadata Model

Portal Label	Element	Is Required
Contributing institution	dcterms:provenance	Required
Title	dcterms:title	Required
Date	dcterms:date	Required
-	dcterms:created	Required
Creator	dcterms:creator	Required
Rights	dcterms:rights	Required
Language	dcterms:language	Required
Congress	dcterms:temporal	Required
Collection title	dcterms:relation	Required
Physical Location	dcterms:isPartOf	Required
Collection finding aid	dcterms:source	Required
Identifier	dcterms:identifier	Required
Preview	edm:preview	Required
Available at	edm:isShownAt	Required
Record type	dcterms:http://purl.org/dc/terms/type	Required
-	dcterms:type	Required
Policy Area	dcterms:subject	Recommended
Topic	dcterms:http://purl.org/dc/terms/subject	Recommended
Names	dcterms:contributor	Recommended
Location represented	dcterms:spatial	Recommended
Extent	dcterms:format	Optional
Publisher	dcterms:publisher	Optional
Description	dcterms:description	Optional

Appendix C: DEIA Consultant Call for Proposals

The American Congress Digital Archives Portal is a collaborative, non-partisan project that makes congressional archives from multiple institutions available online, bringing the history of the People's Branch to the people. The project provides open access to archival materials and descriptive information. In the 2022 project white paper, partners laid out plans for the project that included focusing on the contributions underrepresented people and communities have made in Congress by including archives of women, Hispanic Americans, Black Americans, Asian and Pacific Islander Americans, Indigenous Americans, LGBTQ+ Americans, and Americans with disabilities who have served in Congress.

The project is seeking a consultant to begin this Diversity, Equity, Inclusion and Accessibility (DEIA) work. The consultant will create a set of guidelines and actionable steps to act as a roadmap for how the project can best engage with potential partners, archival collections, and user communities. Partners may be institutions that preserve the collections of Members of Congress who identify as part of an underrepresented group. Archival collections may be created by a Member of an underrepresented group or may contain significant materials related to an underrepresented group. Communities may be scholars, students, or others who traditionally do not use congressional archives but who may, through their use, contribute new perspectives and understanding about public policy, American history, or Congress as an institution.

The consultant should have experience leading DEIA work, a successful record of completing deliverables on deadline, and some familiarity with congressional archives and/or congressional history. The consultancy will begin in January 2023 and meet monthly with project stakeholders. The consultancy does not require travel. The consultant will submit a final report to the project director and partners by June 30, 2023. Compensation is \$1,200. This project is made possible in part by a 2022 award from the Catalyst Fund at LYRASIS.

Appendix D: ITAV Toolkit Checklist

Governance Pre-Work	
https://itav.lyrasis.org/toolkit-governance/	
Mission	Need (stated in proposal but get group consensus)
Vision	Need (stated in proposal but get group consensus)
Strategic Priorities	Need (stated in proposal but get group consensus)
Community Stakeholders	Done - NEH Implementation

Governance Phase 1: Establishing Governance	
https://itav.lyrasis.org/toolkit-governance/#gov-phase-1	
1. Catastrophizing	Need
2. Checklist	Need
3. Models/Process	Need
4. Choose Your Own	Need
5. Document/Implement	Need
6. Case for Change	Not sure
7. Evaluate Comms	Not sure

Technology Phase 1: Laying the Groundwork	
https://itav.lyrasis.org/toolkit-technology/#tech-phase-1	
Tech Stakeholders	Done - NEH Foundations
Tech Skills Inventory	Done - NEH Foundations
Personas and Pathways	Maybe
Landscape Analysis	Need
Buy a Feature	Done (end user survey) - NEH Foundations
Community QA	Need

Documentation Friction	Need (after creating end-user documentation)
------------------------	--

Technology Phase 2:

<https://itav.lyrasis.org/toolkit-technology/#tech-phase-2>

8. Shadow Observations	Done - NEH Foundations
9. Recognition	Not sure
10. Value Propositions...	Done - NEH Foundations/NEH Implementation proposal
11. Building Communities	Need
12. Integration	Need
13. Retire Features	Need

Technology Phase 3:

<https://itav.lyrasis.org/toolkit-technology/#tech-phase-3>

14. Tech Stakeholders	Need
15. Catastrophizing Pt. 1	Need
16. List of Dreams	Need
17. Catastrophizing Pt. 2	Need
18. Tech Skills Inventory	Need
19. End-of-Life Comms	Need
20. Winding Down	Need

Resources Phase 1: Creating Consistency

<https://itav.lyrasis.org/toolkit-resources/#res-phase-1>

1. Landscape Analysis	Need
2. Financial Models	Need
3. Human Resources	Need

4. Value Propositions	Need
5. Org Home Matrix	Need
6. Org Home Requirements	Need

Community Engagement Phase 1: Getting Beyond Initial Stakeholders

<https://itav.lyrasis.org/toolkit-engagement/#ce-phase-1>

1. Your Community	Done
2. Create Personas	Need
3. Plan Outreach Comm.	Need
4. Outreach Charter	Need
5. Stakeholder Matching	Need
6. Communication Plan	Need
7. Evaluate Communication	Need