ARCHIVING CONSTITUENT SERVICES DATA OF THE U.S. CONGRESS

A report of the Society of American Archivists
Congressional Papers Section CSS/CMS Task Force

SOCIETY OF
American
Archivists

Congressional
Papers Section

ARCHIVING CONSTITUENT SERVICES DATA OF THE U.S. CONGRESS

A report of the Society of American Archivists
Congressional Papers Section CSS/CMS Task Force

Danielle Emerling, Chair
West Virginia University Libraries

Hope Bibens
Drake University

Natalie Bond
University of Montana

John Caldwell
University of Delaware Library

Alison White
United States Senate Historical Office

November 2017

SOCIETY OF
American
Archivists

Congressional
Papers Section

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Since the 1978 introduction of computerized systems for managing constituent correspondence in the offices of members of the U.S. Congress, systems have become proprietary and more complex. The data exported from these systems and transferred to collecting repositories presents numerous challenges, and despite recent attempts, archivists lack a systematic way to process, preserve, and provide access to constituent data. Without a concerted effort by the congressional archival community, these important historical sources documenting the representative function of Congress are in danger of disappearing.

The Society of American Archivists Congressional Papers Section formed the CSS/CMS Task Force in August 2016 to study the issue, and this report addresses the ongoing development and use of constituent services data in Congress; the vendors, systems, and functionality; and the obstacles faced by collecting repositories. It provides an overview of research conducted with analog constituent correspondence and the vast potential for research with constituent data. The report recommends a vested advocacy coalition to support management guidelines in member offices, in commercial vendors working with congressional offices, and in collecting repositories, and it calls on the community to develop a technological solution for processing, preserving, and providing access to constituent data that will benefit both large and small repositories.

DEFINITIONS

ASCII file – a text file in which each byte represents one character according to the American Standard Code for Information Interchange (ASCII), a character encoding standard for electronic communication

Archive Format – a data export format that consists of 32 fields which was devised by the Senate Sergeant at Arms for transferring data from Constituent Services Systems in congressional offices to collecting repositories

Automated Indexing System (AIS) – a feature of constituent data systems that allows quick and effective indexing of large quantities of files, usually contained within the DAT file upon export

Casework Files – requests and responses to requests for the member's intervention with a federal agency to solve an individual problem that are often handled by the state/district offices; these are usually managed in CSS/CMS databases

Collecting Repository – used throughout this document to refer to any library, archives, historical society, museum, or other institution that houses congressional papers

Constituent Services Systems (CSS) – automated, proprietary systems used in the U.S. Senate to receive, store, index, and send correspondence with constituents and that increasingly are used for document management and scheduling

Correspondence Management Systems (CMS) – automated, proprietary systems used in the U.S. House of Representatives to receive, store, index, and send correspondence with constituents and that increasingly are used for document management and scheduling

DAT (.dat) file – a generic data file that may contain data in text or binary format

Form Letters – letters from interest groups, often duplicative, these typically constitute the bulk of the incoming mail

Form Letter Library – Paragraphs on specific issues drafted in advance and inserted into outgoing constituent correspondence in response to incoming constituent mail addressing that issue

Issue Mail – non-form letter mail, usually from individual constituents

Personally Identifiable Information (PII) – sensitive personal information, such as a Social Security number

Senate CSS Data Interchange Format (SCDIF) –a data export format originally developed to migrate CSS data between proprietary systems and containing over 200 fields

TAB (.tab) file – text file containing a list of data separated by tabs

INTRODUCTION

One of the core functions of the U.S. Congress is representing the views and needs of constituents. Members of Congress serve as constituents' connection to the federal government and represent constituent concerns and state and local issues while debating legislation of national importance. All members maintain state offices to support this important work, and their success or failure in providing these services can be a large factor at election time. This relationship, crucial to a representative democracy, is well documented in congressional collections in constituent correspondence, such as issue mail, and casework.

Issue mail and casework have traditionally been maintained as voluminous paper files, but an increase in digital communications and the transition to proprietary correspondence management systems in congressional offices has changed the landscape for archivists. Constituent Services Systems (CSS) in the Senate, and Correspondence Management Systems (CMS) in the House of Representatives, continue to develop in complexity, and the data exported from these systems poses significant challenges for collecting repositories seeking to preserve the communications between representatives and constituents.

Constituent correspondence data holds unique opportunities for big data research to uncover patterns and correlations in public opinion, civic engagement, and congressional action. Geographic and demographic data tracked along side specific issues, from health policy reform to foreign relations, holds great potential for data analysis and visualization. Yet no repository on its own has been able to systematically process and provide access to the data in a replicable way.

Without a strategic effort in the congressional archival community to process, preserve, and provide access to constituent correspondence data, at best these opportunities will go unrealized, and at worst, data will be lost to the inherent threats of born-digital materials – deterioration and obsolescence. Inaction may also result in fewer transfers of data from congressional offices to collecting repositories, meaning that an entire body of historical documentation of American democracy is in danger of disappearing.

This report offers a primer on the development, functions, and workflows of CSS and CMS; an overview of issues faced by collecting repositories; and a review of the research possibilities and challenges with constituent correspondence. Finally, the report offers guidelines for the short-term management of exported data in collecting repositories and recommended solutions for improving management and preservation of data in the long term.

BACKGROUND

Senate

Automated systems for handling constituent correspondence were introduced in the Senate in 1978.[1] The first systems were developed in-house and required a personal computer in the office to link to the Senate Computer Center network, a centralized Senate system of magnetic tapes used for a variety of office functions, using two related, but separate databases: the Automated Indexing System (AIS) and the Senate Mail File (SMF). There was a separate subsystem called the Casework System (CWS) used to track casework. Certain reports were supplied to offices using these systems. Correspondence (and casework) could be organized by document number or by assigned topic. It was suggested that routine requests, such as flag requests, be filed separately to enable future disposition.

In 1982, the Senate Committee on Rules and Administration issued a "Dear Colleague" letter authorizing the Sergeant at Arms (SAA), upon request of the member, to make available computer tapes containing index and reference codes to correspondence files to a member's designated archival repository for research purposes. The 1992 edition of *Records Management Handbook for United States Senators and Their Repositories* advised offices that it was "imperative that a staff member maintain current lists of the meanings of codes and groups if used. Outdated lists should be retained for the permanent record."[2]

In the early 1990s, independent correspondence management systems created by outside vendors were developed and approved for use in the Senate. Most of these vendors offered Windows-based, relational databases, and the Senate Archivist advised on how to manage information in these systems. Recommendations included tagging constituent entries by type and filing by group type (special interest groups associated with the correspondent); coding issue mail by subject; and using a controlled vocabulary to assign subjects. Vendors included Capitol Correspond, Quorum Power, Quick Response and Neves (a Macintosh system). The Senate Archivist worked with the members of the Sergeant at Arms' Constituent Correspondence Systems team to determine what fields from the system might be useful for export at the time the member left office. This work, coupled with Senator Sam Nunn's interest in using the electronic information as an index to the hard copy, was the origin of the 32-field Archive Format. (See Appendix A for Archive Format fields).

Generally, the exported information included the form letter library, which usually contained pre-written responses on topics; correspondence records, including casework if selected; the AIS data; incoming email text and additional files attached to correspondence records. The member's office would be responsible for adding additional digitized material, such as old libraries/codes from AIS, hard copy correspondence and casework, and schedule data, which offices generally

---

[1] Naomi Nelson, "Taking a Byte Out of the Senate: Reconsidering the Research Use of Correspondence and Casework Files," in *An American Political Archives Reader*, ed. Karen Dawley Paul, Glenn R. Gray, and L. Rebecca Johnson Melvin (Lanham, Maryland: The Scarecrow Press, Inc., 2009), 235-252.

[2] United States Senate, *Records Management Handbook for United States Senators and Their Repositories,* by Karen Dawley Paul, Washington, D.C.: GPO, 1992.

converted to the Microsoft Outlook format. The SAA coordinated the exports with the vendors and reviewed data for quality.

Offices were advised to save reports provided by CSS system vendors; they could run custom reports at an additional cost. Some reports offered to active offices included: noting the volume and type of outstanding (unfulfilled) mail or casework reports, top code usage, top letter system codes, polling interest, and hot topic reports. It was up to individual offices to take advantage of these statistical reports and include them with permanent records when the office closed. The Senate Archivist recommended that repository archivists proactively determine what statistical reports the office had elected to receive and to ask for them to be generated while the systems were still active.

As members' offices began to express interest in receiving a more complete set of data, the SAA considered additional options, including providing departing members with the Senate CSS Data Interchange Format (SCDIF), originally created by SAA and CSS vendors to migrate data between vendors' proprietary systems. Although this is still not a complete download, SCDIF includes over 200 additional fields.

The SAA had concerns about making changes to the export format because of potential budget limitations. Vendor requirements included data export of only 32 fields as part of the existing contract. It was felt that adding anything to that contract, such as asking for an expanded download of data, might present added and unwelcome costs to office budgets. The determinations as to which format is selected is made by the closing office. Frequently there are concerns about including casework (because of privacy issues) and, being rushed and overwhelmed to close the office on a strict deadline, decisions surrounding exported CSS data can become a low priority.

Another concern in the Senate was the ability to screen a larger export for personal information and for corruption, services provided for the 32-field Archive Format by not including casework. Closing offices requested the export, and the exported data was sent from the vendor to the SAA. The SAA verified that the data requested was present, and then sent the data onto the repository. Repositories often received the verified copy of the data a year or more after the office had closed, often long after the collection had arrived, and lacked information about the chain of custody.

At the end of the 113th (2014) Congress, one Senate office bypassed the SAA and exported the complete set of data, not in the SCDIF format. At the end of the 114th (2016) Congress, several offices requested an SCDIF export, providing archivists with the first opportunity to see what the expanded data set looks like, how long it takes data to get to the repository, and how SCDIF data might be managed in a repository setting. The 114th (2016) Congress was the first in which all Senate offices were able to choose between either the Archive Format or the SCDIF format when requesting an exit conversion of their CSS data.

House of Representatives

The use of vendor supplied Correspondence Management Systems (CMS) in the House dates to the mid-1990s. Prior to that time, House Information Resources (HIR) maintained an in-house database called the Member Information Network (MIN). This was not a true CMS like those in use today, but functioned more as a precursor to the Legislative Information System (LIS) and THOMAS databases. During this period, private database vendors began offering their services to House members using the MIN network. These vendors were required, at request of the Member, to provide their entire databases on 9-track ASCII formatted tapes.

Archival repositories receiving these materials were advised to request a copy of the system operating manual and a history of how they were coded (e.g., by topic, document type, revisions) with the expectation that staff turnover might lead to various versions of coding systems.[3] Micro MIN, an in-house CMS database created and managed by HIR, followed MIN. In 2000, Micro MIN disappeared to make way for private sector vendors as the exclusive providers of CMS software and services.

CMS systems developed so that in addition to the primary function of managing incoming and outgoing correspondence and casework, some vendors bundled e-newsletter capabilities, social media integration, and legislation status tracking that offers constituents the opportunity to comment on legislation while it's pending. Most vendors offer a method for extracting and transferring data between CMS products when the member is continuing in elected office either in the House or when moving to the Senate.

Under the leadership of Karen Haas, Clerk of the House, the Advisory Committee on the Records of Congress addressed this issue in the 113th Congress. Initially, the House Interchange Standard (HIS) was developed to exchange data between vendors, but one vendor developed an alternative Microsoft Access database form that is now offered by all vendors.[4] Senate collections are unable to take advantage of the Access database because the volume of data exported from Senate constituent correspondence systems generally exceeds the capacity of the Access database.

Members who are leaving office choose to export their CMS data (either a subset or in its entirety) on a case-by-case basis. At this time, the House does not maintain statistics on how many members take CMS data when they leave office.

---

[3] Joni Bell, "Congressional Demographics – It's [sic] Effect on Member of Congress Files," Society of American Archivists Congressional Papers Roundtable newsletter, December 1987, 5-7, https://www2.archivists.org/sites/all/files/1987_December.pdf.
[4] Advisory Committee on the Records of Congress, meeting 46 minutes, June 16, 2014, https://www.archives.gov/files/legislative/cla/advisory-committee/minutes/2014-19-06.pdf.

VENDOR INFORMATION

The Senate and House of Representatives share similar vendors and products, but rules for procuring and managing systems are unique to each chamber. As of 2016, four companies supported CSS systems in the Senate: iConstituent, Leidos Digital Solutions (formerly Lockheed Martin), ComputerWorks and Symplicity. Of the four vendors, Leidos' product Intranet Quorum, better known as IQ, is the most widely-used system. Six CMS vendors with approved contracts provide CMS services in the House: Leidos Digital Solutions, iConstituent, Fireside21, ComputerWorks, HouseCall IT, and Symplicity. Leidos is the predominant vendor in the House. As of 2014, more than half of Member offices had contracts with Leidos.

CSS vendors - U.S. Senate	CMS vendors - U.S. House of Representatives
1. iConstituent 2. Leidos Digital Solutions (formerly Lockheed Martin) - Intranet Quorum (IQ) 3. ComputerWorks - InterTrac 4. Symplicity - Voice	1. iConstituent 2. Leidos Digital Solutions – Intranet Quorum 3. ComputerWorks - InterTrac 4. Simplicity - Voice 5. Fireside21 6. HouseCall IT

Figure 1: CSS and CMS vendors and products as of 2016

In the Senate, CSS contracts are negotiated on a six-year cycle by the SAA, which also manages the server space on which CSS data is stored in the Senate complex. Funding allocations for a CSS instance are based on state population. During the transition period for new members, (Election Day to January 3), vendors give live demos to senior staff who will select a system and stand it up prior to the start of the new Congress. If a new Senator has existing CMS data from a House CMS, that data will be transferred into their new Senate system.[5]

State Population	Allocation (per 6-year term)
≤ 8 million	$400,000
> 8 million	$450,000
New members (regardless of population)	$15,000 (one time supplement)

Figure 2: U.S. Senate funding allocations for CSS as of 2016

At this time, there is no public information regarding CMS vendor contracts or CMS funding allocations for the House of Representatives.

---

[5] United States Senate, Office of the Sergeant at Arms, CSS III Solicitation 20160R-025, 2016.

CSS/CMS FUNCTIONALITY AND DATA LIFECYCLE

The CSS/CMS have two essential functions: communications management and workflow management. Specifically, the systems support processes related to tracking, responding to, and managing incoming and outgoing communications between constituents, members, and government agencies. They also serve as a data repository, allowing offices to repurpose data collected from constituents communicating with the office. This data includes (but is not limited to) names, contact information, demographic information, geographic information, and, possibly Personal Identifiable Information (PII). While they are not explicitly constructed to serve as a document management tool, systems have the functionality to store and index records and provide a means to search the repository and recover records. Vendors also offer a scheduling module to manage the calendars and personal schedules of members and staff.

Offices primarily use systems to track and store incoming and outgoing correspondence with constituents, and many members also use systems to manage casework. Correspondence can be submitted via email, through the member's website as an online form, or through physical mail. Physical mail received by the office could either be scanned into the CSS/CMS, or a synopsis of the letter could be entered into the system manually. These systems can also be used to track incoming constituent calls on policy issues and manage the member's schedule. For scheduling, however, many offices opt to use Microsoft Outlook, particularly for the D.C. schedule. It is possible to migrate scheduling data between the two utilities.
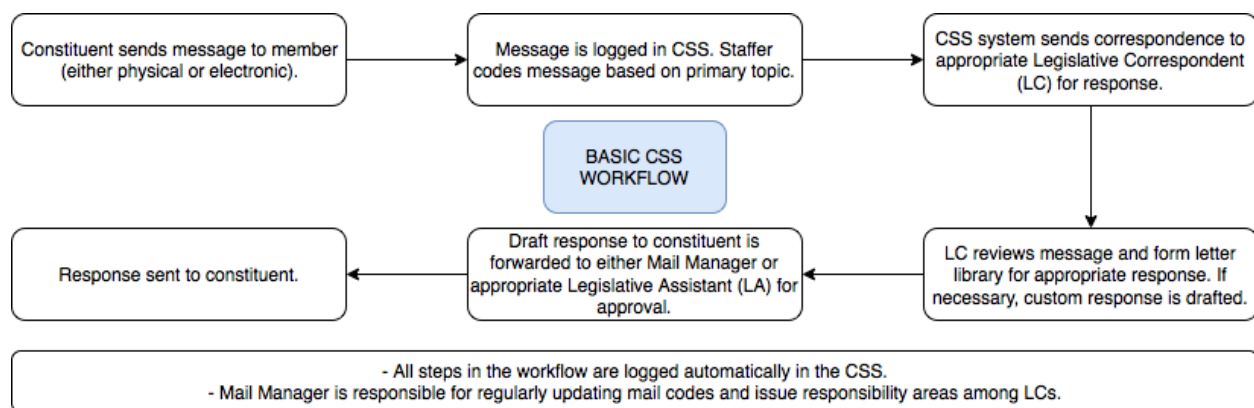
Figure 3: Basic CSS workflow

Because CSS/CMS instances have document management functionality, some offices attempt to manage documents other than correspondence and casework records in the systems. For example, IQ now offers an optional module with a document hierarchy for final versions of documents and export capabilities for certain file formats. Unfortunately, there is limited metadata preservation within the current module, unless the metadata is included within the document (e.g., dates, names) or document naming convention (e.g., "Mmo" for memo or "Corr" for correspondence). Some offices are implementing other content management systems, such as Microsoft SharePoint, to track constituent responses to pending legislation and "hot topics" (top policy issues or events prompting constituent contact), and generate corresponding reports. These systems are duplicating the functionality, but not necessarily the content, of the CSS/CMS.

Additional uses for CSS/CMS in active offices include tracking grant programs and awards to the state, constituent outreach and services (e.g., flag requests, tour requests), e-newsletters, and issue and vote tracking. There are also a variety of reporting options and frequencies available to staff. (See Appendix B for a list of reporting options).

Vendors are able to extend the functionality of their system with optional software modules (e.g. the document management module) for specific office needs, so long as the module meets the minimum software requirements in the contract and can be integrated with the existing functionality of the CSS. Further, this additional development work must have a demonstrable additional cost to the vendor and meet a need servicing ONLY a single or select group of offices. Changes to CSS are reviewed by SAA to ensure that new modules conform to existing export requirements and data extraction complies with SCDIF; if the new module is not SCDIF compliant, the vendor must provide an updated SCDIF to SAA for approval.

System functions used in each office, as well as contractual agreements, have serious archival implications. These include issues of data integrity, accuracy, and completeness, such as what, if any, data manipulation occurred; custodial history; and availability of necessary metadata documenting provenance, technical, and administrative information.

Data Ownership, Backups, and Retention

All CSS/CMS data is considered the property of the member, a fact reflected in vendor contracts. As a result, system data is stored on virtual servers located on the Capitol campus, and members reserve the right to keep, move, or export the entirety of the data at their discretion. However, not all modules are designed to accommodate archival best practices (e.g., maintenance of technical metadata associated with documents stored in the system).

Sometimes members choose to move, or "gift," data to an incoming member, and this can be a complete data transfer between members of a state delegation or, in some instances, may take the form of open casework or unanswered correspondence. In these cases, establishing provenance can be problematic. Actions may not be documented, except *within* the correspondence documents, listing the dates of interactions or addressing certain staff which indicate the original individuals creating and managing this information.

Vendors are responsible for configuring regular backups of data and ensuring that all data and applications are fully backed up. SAA provides the storage infrastructure for CSS, ensuring that Senate data remains in the physical custody of the Senate. Additionally, vendors must preserve the confidentiality of all information and communication they have access to as CSS data is considered sensitive and is replete with PII.

No public information is available for how the House of Representatives manages CMS and data backups.

Transfers and Exports in the Senate

Data may be converted for the purpose of transfer between member offices or between offices and collecting repositories. A member's office may ask a vendor, via the SAA, to transfer data from the House of Representatives. Numerous offices have noted difficulties in transitioning between systems, including lost data and improperly imported or parsed data, leading to difficulties in searching older data originally created in a different system. These issues of authenticity may be further complicated if data is transferred between members (e.g., open casework).

Standard workflow types that are part of a CSS installation must have full documentation logged with and approved by the SAA. This workflow data is included in the SCDIF. If an office switches vendors, new vendors have five days to upload received SCDIF data into the new CSS database.

For exit conversions, documentation is submitted to the SAA three weeks before the start of the data conversion. Vendors have up to nine days to convert and validate data for export. Offices may request a conversion of scheduling data into Outlook in addition to including data in SCDIF (there is no scheduling data included in the Archive Format). Vendors and the SAA are required to retain a copy of every office's CSS data for one year following an exit conversion. Data are stored offline in both the vendor's native CSS format as well as the conversion format.

Vendors perform the conversion on an SAA provided virtual server on the Senate network; vendors cannot move or copy data from the Senate network. SAA provides a separate Virtual Machine Image (VMI) environment onto which the vendor copies the data, does minimal, if any, data cleanup and converts data into either the archive or SCDIF format. The vendor then copies the data to external storage media and sends this to the SAA. The vendor is responsible for validating the data after copying from the server and again after the data conversion. Vendors are required to provide the member's data in both the proprietary format and the archive/SCDIF format, depending on which format is requested, to the SAA. The SAA will transmit data either to the archival repository (this is known as an exit conversion) or to the new vendor for migration into their system. There is no set timeframe for SAA to send CSS data to repositories. When data is exported in the Archive Format, removal of PII can be handled by not including casework in the export. Casework cannot be excluded in the SCDIF format, and offices are advised to address this issue directly with the receiving repository.

As part of the conversion vendors must provide to the SAA the CSS data libraries, which include a full description of the database schema, data constraints, triggered activities, index information and data access processes, and dependencies. These libraries are not included with documentation sent with the CSS data from the SAA to collecting repositories, to our knowledge. The only accompanying documentation sent by SAA describes the fields in the archive or SCDIF format.

COLLECTING REPOSITORY LANDSCAPE

Collecting repositories must work with congressional offices to obtain CSS/CMS data. In the Senate, a member must formally request the download through the SAA. Vendors may not make the data available directly to a repository. In the House, a member may request a download, or the repository may purchase a download from the vendor with the member's permission. Repositories receiving Senate data in the Archive Format will receive .TAB or .DAT files, accompanied by a separate correspondence library file. Correspondence is typically received as text (.txt), Microsoft Word (.doc), or Portable Document Format (.pdf) files.

The congressional archives community has attempted to address the difficulties associated with processing CSS/CMS data for nearly three decades. With the growth of computer use in congressional offices in the 1990s and early 2000s, the prospect of a substantial electronic component within any collection increased with the close of each Congress. The Society of American Archivists Congressional Papers Roundtable (CPR), renamed the Congressional Papers Section in 2016, addressed issues related to the collection of constituent data and casework at various meetings as reported in CPR newsletters as early as the 1980s.

A session titled "Computers on Capitol Hill" presented in 1987 described the distributed aspects of early House and Senate systems.[6] In 1994, presentations included "Constituent Correspondence and CMS" and "From 'ROBO' Letters to E-mail: The Evolution of the Constituent Mail Function in the Senate" by Naomi Nelson.[7] In 1998, the CPR meeting minutes included a checklist for CMS files created by Nelson with recommended data specifications for receiving these electronic records. These included the recommendation that fixed field data be sent in ASCII.[8] As interest in digital preservation of electronic records developed in the congressional archives community, collecting repositories focused increasingly on ways to properly accept and handle electronic records.

Recent Attempts to Process Constituent Data

Several repositories have attempted to process constituent data and make it available to researchers, but technical and fiscal resources have left most projects unfinished and unscalable. In 2014, the Albert Gore Research Center at Middle Tennessee State University (MTSU) undertook a project to reconstruct the IQ databases for both the Congressman Bart Gordon Papers and the Congressman Zach Wamp Papers, developing a database search engine. The program developed by MTSU allows users to choose a type of data (correspondence or casework), search by correspondence topic, city, county, or ZIP code, and specify a date range. The database is only accessible to individuals with an MTSU ID from a wired campus

---

[6] "Computers on Capitol Hill: SAA in NY," Society of American Archivists Congressional Papers Roundtable newsletter, December 1987, 4, https://www2.archivists.org/sites/all/files/1987_December.pdf.

[7] Naomi Nelson, "From 'ROBO' Letters to E-mail: The Evolution of the Constituent Mail Function in the Senate," Society of American Archivists Congressional Papers Roundtable newsletter, April 1995, 3-5, https://www2.archivists.org/sites/all/files/1995_April.pdf.

[8] "Congressional Papers Roundtable Minutes 1998 Annual Meeting Orlando, September 4," Society of American Archivists Congressional Papers Roundtable newsletter, November 1998, 1-4, https://www2.archivists.org/sites/all/files/1998_November.pdf.

computer on a case-by-case basis. The Bart Gordon Papers finding aid notes that access to the Intranet Quorum database is based on availability of qualified staff and technical resources to conduct queries.[9]

The Robert C. Byrd Center for Congressional History and Education at Shepherd University also embarked on a project to make the systems data of Senator Byrd accessible in 2014. Byrd Center staff worked with a contractor to develop an SQL database, using Microsoft Access as a front-end. The custom SQL database can ingest data, parse the coded fields, and populate forms with those fields in order to provide the researcher with a copy that resembles the original. Due to extensive variations and errors in the file/issue codes generated when the records were created, as well as security concerns, the project was not completed utilizing this initial approach.

That same year, the University of Montana's Mansfield Library Archives and Special Collections began managing systems data exported from the office of Senator Max Baucus. The Baucus data arrived in two batches, one a .DAT file that contained AIS data from 1979-1990, and the other a .TAB file that arrived on an external hard drive that contained IQ and Voice data from 1983-2014. Similar to the Byrd Center, Archives and Special Collections staff imported the data into a Microsoft Access database, where they were able to view the data by field. Significant import errors, however, rendered the dataset incomplete, and no further efforts have been made to review or provide access to the data.[10]

In 2016, West Virginia University Libraries began developing an open source system using a MySQL database that was able to access and search a portion of the constituent data from Senator Jay Rockefeller's office. The system also conducted a geographical analysis and visualization of the data. Development is ongoing, and progress updates are available on GitHub.[11]

Survey Results and Common Obstacles

Based on surveys conducted by the Congressional Papers Section in 2009 and 2014, it is estimated that 15-21 repositories currently hold CSS/CMS data. Repositories that retain systems data have reported receiving correspondence, casework, attachments in their native form, and other types of electronic records. Subsequent levels of access and preservation are determined by the repository's technical capabilities and resources, and on the whole, repositories have reported significant challenges in accessing and processing CSS data. Further, issues of quality and consistency in data exports have been reported. Common obstacles include:
- Proprietary systems are too expensive to purchase.
- Data exports are too large and complex for processing with Microsoft Access databases (often the most accessible, and sometimes only, database software available to repositories).
- Data cannot be reviewed, arranged, or described if the repository cannot open the files.

---

[9] "Bart Gordon Papers," Albert Gore Research Center, Middle Tennessee State University, Murfreesboro, TN, 2013, http://gorecenter.mtsu.edu/research/BartGordon.shtml.

[10] Natalie Bond, "Testing the Waters: Working with CSS Data in Congressional Collections," *Society of American Archivists Congressional Papers Roundtable Electronic Records Case Studies Series*, August 2015, https://cprerc.wordpress.com/case-studies/case-study-004/.

[11] "Rockefeller-CSS," West Virginia University Libraries, 2016, https://github.com/wvulibraries/rockefeller-css.

- Data is sometimes incomplete or missing.
- Records may be duplicated within the data export.
- Correspondence may be missing.
- Exports may contain an inconsistent number of fields (across years and offices).
- Identifiers and metadata may be incorrectly linked.
- Data contains PII that must be reviewed and redacted, preferably in an automated way, before access can be granted.
- Subjects are coded inconsistently within and across congressional offices.

Technology does not always pose the only obstacle, however. Communication between congressional offices and repositories prior to deposit can vary, and the potential exists for large delays, sometimes more than a year, between the time an office closes and the time that a repository receives a final data export. Additionally, congressional staffers quickly migrate to new jobs once an office closes, making it difficult for archivists seeking contextual information about the data to establish lines of communication with those with direct knowledge of how it was originally collected, stored and utilized.

These factors can prevent archivists from gathering illuminating information regarding automated backups of the CSS/CMS, codified office procedures for collecting and entering constituent data, in-office customizations of systems, and other details which could inform archival preservation and processing of data. More generally, they have the potential to dilute repositories' overall understanding of how constituent systems were utilized on a day-to-day basis in their respective offices.[12]

In February 2017, members of the CSS/CMS Task Force reached out to repositories that house data in their congressional collections in an effort to gather qualitative information on work that has been done with systems data, as well as any obstacles encountered. The Task Force received responses from three institutions that hold at least eleven sets of systems data, all of which reflect both the complex nature of the data and the varying technological capacities of archival repositories to ingest, view, manipulate, and preserve the data.

Several trends reflective of the 2014 survey results have begun to emerge from taking a broad look at both these responses and the 2014 survey, highlighting the need for the standardization of systems data exports, greater transparency and cooperation from software vendors, and an open-source database capable of ingesting and aggregating the data. (See Appendix D for more detailed survey responses).

---

[12] Adriane Hanson, "Donations of Digital Records from Congressional Offices: Lessons Learned from the 2014 Election Cycle," *Society of American Archivists Congressional Papers Roundtable Electronic Records Case Studies Series,* June 2015, https://cprerc.wordpress.com/case-studies/donations-of-digital-records-from-congressional-offices-june-2015/.

RESEARCH POTENTIAL

The potential for use of data from CSS and CMS products extends across disciplines, such as history, political science, and public administration, and beyond the academy, to public interest groups and journalists.

Constituent files provide unique opportunities to broadly explore the perspectives of everyday Americans about the legal, economic, and cultural issues affecting their lives. Constituent correspondence, or issue mail, and case files offer windows into public opinion in a state or district, and, when taken with other sources such as voting records, enable researchers to explore the influence of public opinion on elected officials. Case files vividly illustrate the representative function of an office, particularly to marginalized groups whose voices are often absent from the historical record. These sources are one of the few places to observe and compare the responses from members of Congress to individuals and businesses in home states, districts, and across the nation.

Scholars have used constituent correspondence for numerous inquiries, such as exploring how the southern public responded to the Vietnam War,[13] how congressional leaders navigated the complex politics and district dynamics involved with the push for racial equality in the 1950s and 1960s,[14] and how the public's views relating to the Iran-Contra investigation challenged the media discourse about events and expressed "sophisticated visions of democracy, heroism, and patriotism."[15] They have also used correspondence to help explain how Americans reflected upon overseas commitments in the post-Vietnam War era;[16] how constituents aided the war effort in World War II in unique ways, reporting evidence of waste and corruption among war contractors located in their communities;[17] and how individual letters to members of the Senate helped to shape committee investigation and related policy.[18]

However, research such as this is currently conducted with analog (paper or microfilm) correspondence, and is incredibly time and resource intensive. Researchers must identify collections that may contain the information they need and often travel long distances to spend time in the archives. Once there, analog materials need to be hand-coded to quantify information. Another discovery, often made after arriving to do research, is that the entire body of constituent correspondence has not been saved; rather, a sampling of the correspondence is made available to the researcher. It is not always clear to the researcher if they are working from a sample or not,

---

[13] Joseph A. Fry, *The American South and the Vietnam War* (Lexington: The University of Kentucky Press, 2015).

[14] Neal Allen, "Research in Congressional Holdings," presentation to the Association of Centers for the Study of Congress 12th annual meeting, Washington, DC, May 13-15, 2015.

[15] Amy Fried, "Public Opinion in the Archives," in *Doing Archival Research in Political Science*, ed. by Scott A. Frisch, Douglas B. Harris, Sean Q. Kelly, and David C.W. Parker (Amherst, NY: Cambria Press, 2012) and David Thelan, *Becoming Citizens in the Age of Television: How Americans Challenged the Media and Seized Political Initiative During the Iran-Contra Debate* (Chicago: The University of Chicago Press, 1996).

[16] Natasha Zaretsky, "Restraint or Retreat? The Debate over the Panama Canal Treaties and U.S. Nationalism after Vietnam," *Diplomatic History*, 35, no. 3 (June 2011).

[17] Katherine Scott, "A Safety Valve: The Truman Committee's Oversight during World War II," in *Congress and Civil-Military Relations*, ed. by Colton C. Campbell and David P. Auerswald (Washington, D.C.: Georgetown University Press, 2015).

[18] Michael Perino, *The Hellhound of Wall Street: How Ferdinand Pecora's Investigation of the Great Crash Forever Changed American Finance* (New York: The Penguin Press, 2010).

as there is not always information provided about the basis of the sampling and when or why it was done a certain way. This can pose significant challenges to the defense of data-driven conclusions if all the data is not available for use or if there is no clear documentation as to why or how certain data was selected.

Making the digital data from CSS and CMS systems available has the potential to eliminate these steps, and access to complete digital data would enable greater confidence in a researcher's conclusions. Providing online access to these data offers a larger group of researchers the chance to glean potential patterns, trends, and associations made possible by big data sets. The aggregate data opens research potential for analysis, visualizations, digital humanities projects, and more. Correspondence on a range of topics, coupled with demographic and geographic information, presents new opportunities to investigate public opinion and influence over time.

GUIDELINES FOR COLLECTING REPOSITORIES

Given the ongoing challenges associated with CSS/CMS data, specifically the lack of procedures and tools to process and provide access to data, the Task Force recommends that collecting repositories take the following steps to ensure data are better transferred and preserved in the short-term. Communicating with the congressional office, asking for additional information beyond the basic export, and following best practices for managing born-digital materials are the most important actions a repository can take.

1. The archivist should contact the congressional office to express interest in archiving the system data as soon as the repository is chosen. The Chief of Staff or the systems administrator are the individuals most likely to have knowledge of the system and/or contact with the vendor.

2. In addition to the data, the archivist may wish to request the following:
- All system documentation, including the system name and version(s), mail and casework manuals, and dates of updates
- Legislative correspondent/legislative assistant Issue Lists (including historic lists)
- System staff (list of active and inactive staff members)
- Workflow processes and workflow summary
- While the data is within the live system in the office, the archivist may also request copies of any reports that are routinely run for the office and any custom reports that the repository may identify as useful or desirable. These might include:
  - Hot topics reports
  - Week in Review
  - Mail volume by topic report
  - Code usage – Interest code, Letter code, or Personal code usage in given period
  - Topic codes and subcodes by year
  - Correspondence Detail Report by Group (snapshot of correspondence based on specific group and date range)
  - Casework reports, especially if the office is NOT transferring casework to the repository

3. The archivist should contact the Mail Manager, Correspondence Manager, or other individual in charge of managing the daily correspondence in the office to learn how the system was managed. Helpful questions include:
- How was the system used in the office?
- How was correspondence coded?
- Are retired codes saved?
- Who was responsible for coding incoming correspondence over time?
- Were phone calls entered into the system?
- What other schedule or document management system, such as Microsoft Outlook or SharePoint, is in place, and is it integrated with the CSS/CMS?
- What reports did the office generate from the system?
- What training is provided to staff using the system?
- Are there unique issue codes or patterns used by staff *at any time*?
- Did the office use multiple systems? If so, when did the office transition systems?

4. If the House office did not choose to export system data upon closing the office, the repository may contact the vendor directly with permission from the member. Repositories should request a full data export and expect to receive the data on external media, such as a CD, flash drive, or hard drive. Additionally, repositories should request all system documentation, including information about fields and functions.

5. Generally, any data export will include the form letter library, correspondence records (and casework if selected), the AIS in the DAT file, and any incoming email text or additional attachments.

6. When the export arrives, follow best practices for born-digital materials to accession the data:
- Use a "clean" computer and a write blocker to safely copy data.
- Copy files using the repository's preservation system or free utility, such as Bagger, to secure storage where files receive regular backup.
- Generate and record a checksum and file format identification.
- Verify that the data has been exported correctly.
- Document all actions taken where other collection information is stored (e.g., accession record, finding aid, etc.).
- Return the external media to storage and retain in the short-term.

RECOMMENDATIONS FOR IMPROVING THE LONG-TERM PRESERVATION OF CSS/CMS DATA

A review of the history of the management of CSS/CMS data in both chambers of Congress and in collecting repositories reveals a category of materials at high risk of disappearing from the historical record. The solution is twofold. First, Congressional offices must appreciate the historical importance of archiving constituent correspondence data and manage it in a way that facilitates archiving. Second, collecting repositories must develop a standardized way to systematically and automatically process and preserve the data, redact PII, and securely provide access to this rich historical source, which if successfully uploaded to a purpose built database would be even more valuable than its former paper based form.

Interest and support from members of Congress in the archiving endeavor, as well as thoughtful and incremental change, have historically led to better outcomes. Therefore, this Task Force strongly urges building a coalition to advocate for the recommendations contained in this report. This coalition should include current and former members of Congress, the Congressional History Caucus, the National Coalition for History, the Congressional Management Foundation, the National Association of Government Archives and Records, the Council of State Archivists, the SAA Congressional Papers Section, and the Association of Centers for the Study of Congress.

Congressional Office Data Management
- Introduce authority control to standardize subject coding across systems and offices to ensure data are interoperable and searchable across collecting repositories
- Provide better training to staff to more consistently apply subject codes to increase in-office efficiency and interoperability in the archives
- Encourage staff to consistently run and archive recommended reports
- Maintain all mail issue assignments for Legislative Assistants and Legislative Correspondents over time
- Maintain all prior subject code lists and transfer lists to offices and archives inheriting system data

Collecting Repository Data Archiving
- Contact offices as soon as possible to express interest in receiving the data and reports
- Advocate for the fullest export possible
- Request reports of any data, such as casework, that is not transferred
- Maintain a record of research use with constituent correspondence to help strengthen the case for preservation
- Work as a community to develop an open access tool to process and provide access to the data using one of the following models:
    - Identify a single collecting repository that will secure funding for development, develop a tool, make it available to the community, and maintain the system
    - Work within an existing organization to secure funding for development, develop a tool, make it available to the community, and maintain the system
    - Create a new consortium to secure funding for development, develop a tool, make it available to the community, and maintain the system

APPENDIX A: THE ARCHIVE FORMAT

Field No.	Field Name	Description
1	Prefix	Prefix (ex. Mr. or Mrs.)
2	First name	First name
3	Middle name	Middle name or initial
4	Last name	Last name
5	Suffix	Suffix
6	Appellation	Appellation (ex. MD)
7	Title	Org. Title (ex. President)
8	Organization	Organization or business name
9	Address line 1	First line of address
10	Address line 2	Second line of address
11	Address line 3	Third line of address
12	Address line 4	Fourth line of address
13	City	City
14	State Code	2-character state code
15	Zip Code	Zip code for domestic address
16	Country	Country name (internat'l address)
17	In Correspondence ID	Numeric identifier of incoming corr
18	In Correspondence Type	Type of incoming corr – CASE or GEN
19	In Correspondence Method	Method of contact (Letter, Email, Phone)
20	In Correspondence Date	Date Incoming corr received (YYYYMMDD)
21	In Correspondence Topic(s)	Topics that describe corr (ie War on Terror)
22	In Correspondence Text	Freeform text w/ notes or comments
23	In Correspondence Doc Name(s)	Name/path of corr files/email attachments
24	In Correspondence Filing #	Identifier providing location of hardcopy doc
25	Out Correspondence ID	Numeric identifier of outgoing corr
26	Out Correspondence Type	Type of outgoing corr - CASE or GENERAL
27	Out Correspondence Method	Method of contact (Letter, Email, Phone)
28	Out Correspondence Date	Date Outgoing corr closed (YYYYMMDD)
29	Out Correspondence Topic(s)	Topics that describe corr (ie War on Terror)
30	Out Correspondence Text	Freeform text w/ notes or comments
31	Out Correspondence Doc Name(s)	Name/path of out corr files/email attachments
32	Out Correspondence Fill-in Fields	Name of fill-in field as it appears in doc followed by PIPE and text inserted into doc in place of fill-in field

APPENDIX B: SAMPLE REPORTS

The following is a list of the reports that can be generated from CSS/CMS and exported to CSV. These reports can be generated at varying degrees of frequency (daily, monthly, quarterly, yearly, by Congress, etc.) on either an ongoing basis or before an exit conversion. Exact report names are based on a 2016 iConstituent instance.

- Casework: A summary of all cases. Includes Casework ID, case type, received, created and closed dates, status, the staffer assigned to the case, county and zip code information, along with any notes or comments added by the staffer working the case. Some PII may be present, particularly in the Zip (with the 5-4 zip code information) and in the notes fields.
- Casework Production Stats: Lists the type of case, number of cases received, number of outstanding cases and closed cases.
- Casework Type: Lists all of the casework codes with a description of the type of case, as well as the number of open, closed, and total cases based on type code.
- Correspondence: This report has three tabs: Correspondence, By County and By County and Interest.
  - Correspondence: This is a list of all correspondence received by interest code.
  - By County: Lists the total number of correspondences received from each county.
  - By County and Interest: Total number of correspondence by interest code received from each county.
- Interest Codes: List of all of the casework and letter interest codes created in the CSS system.
- Letter System Codes: List of all letter codes with a description of the issue, the creator of the particular code, and the number of letters received based on each code.
- Mail Code Usage: There are three types of Mail Code Usage reports included: Group, Interest and Letter.
  - Mail Code Usage Interest: Records the number of incoming correspondence assigned a particular interest code when the piece of mail was entered into the system.
  - Mail Code Usage Group: Records the number of incoming correspondence assigned particular group code when the piece of mail was entered into the system. Group codes may be used as a second level of identification about the major topic of the letter (e.g. Health-- Affordable Care Act). This is meant to be a subset of the interest code that is more specific than the interest code assigned to the letter when the correspondence is entered into the system.
  - Mail Code Usage Letter: Includes the code, a description of the issue relating to the code assigned to the correspondence when entered into the system and the number of times that code was used over a period of time.
- Personal Codes: Personal codes are assigned to constituent contacts, not letters, and are typically used to describe something about the constituent.
- Staff: List of all staff members with access to the CSS system. May contain some personal information about particular staffers.

APPENDIX C: SENATE CSS DATA INTERCHANGE FORMAT (SCDIF) FOR DATA ARCHIVING

General Requirements

This document defines a common format for the information contained in Constituent Services Systems (CSS). This format will be used to transfer data from one CSS to another, and may also be used to transfer data from an outgoing Senator's CSS system to their designated archival repository (SCDIF archiving conversion).

The data will be provided on a USB hard drive. Records will be variable length ASCII with carriage return line feed (CRLF) at the end of each record. TAB characters will separate fields. Records will be filled out with extra TAB characters at the end so as to have the same number of TAB characters (fields) in each record. It is the responsibility of the vendor who creates the data to make sure that there are no TAB characters or CRLF in the data fields themselves. Any CRLF within data fields shall be replaced with a PIPE character (|). Fields may be any length but no record can be longer than 32K.

Information in the database will be classified into categories:

1	Person data
2	Communication data
3	Workflow data
5	Household data
6	Form Letter/Document Library data
7	Schedule data
8	Code tables

Each record on the CSS data file will have a 2-byte Record Type field with the first byte being a number identifying one of the categories above and the second byte being a letter identifying the type of data within that category.

SCDIF archiving conversions will be customizable to the extent that the Member offices will have the option to include or exclude all or some of the category 3-workflow data. Offices will also have the option to include or exclude all of the category 7-schedule data.

Person Data

1A Record – Name Data

Field No	Field Name	Comments
1	Record Type	"1A" - indicates that the record contains person name information
2	Person ID	A unique ID which identifies the person.
3	Person Type	A code that identifies the type of person in this record. Use the appropriate entry from the table below.
4	Prefix	Prefix of person such as "Mr.", "Mrs." or "The Honorable".
5	First Name	First name of person.
6	Middle Name	Middle name or initial of person.
7	Last Name	Last name of person.
8	Suffix	Suffix of person such as "Jr.".
9	Appellation	Appellation of person such as "MD".
10	Organization Name	Name of the organization. This field is used if the record is not for an individual.
11	Salutation	Preferred salutation used for the person, such as "Bob" or "Senator Jones".
12	Date of Birth	The person's date of birth in the format YYYYMMDD.
13	No Mail Flag	"Y" if the person has requested that no mail be sent to then at any of their addresses.
14	Deceased Flag	"Y" if the person is deceased.
15	Spouse's Name	Name of the person's spouse.
16	Email Flag	"Y" if the person has requested that their preferred method of communication is via email.

Person Type Values

Value	Meaning
AG	Agency
CS	Constituent
MC	Member of Congress
OR	Organization
PRESS	Press
USER	Internal Staff

1B Record – Address Data

There may be multiple 1B records for a person, one for each address.

Field No	Field Name	Comments
1	Record Type	"1B" - indicates that the record contains person address information.
2	Person ID	ID which identifies the person. This ID field must be the same as the ID field in the associated 1A record for the person.
3	Address ID	Numeric ID or sequence number which identifies the address.
4	Address Type	The type of address. Enter the appropriate code from the table below.
5	Primary Flag	"Y" if this is the primary address for this address type, or the only address of this type. For example, if the address type is "BU" and this flag is "Y", then this is the primary business address.
6	Default Address Flag	"Y" if this address is the default mailing address for the person.
7	Title	Title associated with an organization or business address. For example, "CEO" or "President".
8	Organization Name	Name of an organization or business such as "American Medical Association". This name is used if the person record is for an individual and this is his/her business address.
9	Address line 1	First line of address.
10	Address line 2	Second line of address.
11	Address line 3	Third line of address.
12	Address line 4	Fourth line of address.
13	City	City
14	State Code	2 letter state code such as "VA", "DC" or "MD".
15	Zip Code	9 digit zip code. Format for 9 digit zip code is XXXXX-XXXX. If only the 5-digit code is known, then the format is XXXXX. If international address, enter the appropriate zip code of the country.
16	Carrier Route	Carrier route.
17	County	County.
18	Country	If international address, enter the country.
19	District	Congressional district.
20	Precinct	Precinct.
21	No Mail Flag	"Y" if no mail should be sent to this address.
22	Deliverability	"U" if this address has been marked as undeliverable, "D" if marked as deliverable, "P" if marked as possibly deliverable, and NULL if address deliverability has not been verified. (Note: please see attachment XXX.)

Address Type Values

Additional "address types" may be added as needed, but the vendor must document all non-standard types that have been added.

Value	Meaning
BU	Business

Value	Meaning
HO	Home
IN	International/Foreign

1C Record – Person Codes

There may be multiple 1C records for a person.

Field No	Field Name	Comments
1	Record Type	"1C" - indicates that the record contains person code information.
2	Person ID	ID which identifies the person. This ID field should be the same as the ID field in the associated 1A record for the person.
3	Person Code Type	Type of code entered in the following field. Enter the appropriate code type from the table below. The same code type may be used more than once.
4	Code	Code to describe the person/agency. There must be an 8A record containing this person or workflow code.

Person Code Type Values

Value	Meaning
PERS	Code to describe the person (e.g. DOCTOR, MAYOR, VIP, POI).
WORK	For agency records, code to describe the agency/workflow (e.g. SSA).

1D Record – Person Text (Comments or Notes)

There may be multiple 1D records for a person.

Field No	Field Name	Comments
1	Record Type	"1D" - indicates that the record contains comments or notes about the person.
2	Person ID	ID which identifies the person. This ID field must be the same as the ID field in the associated 1A record for the person.
3	1D Sequence Number	Sequence number that identifies the order of 1D records for a person. A sequence number must be provided if fields 5/Date and 6/Time are NULL.
4	Text	Freeform text containing comments or notes about the person.
5	Date	Date the text was added to the record. Must be in YYYYMMDD format.

Field No	Field Name	Comments
6	Time	Time the text was added to the record. Must be in HH:MM:SS format (e.g., 2:20 PM will be provided as 14:20:00).
7	User ID	User ID of the staffer who entered the text. There must be an 8A record containing this staff user ID.

1E Record – Person Phone/Email

There may be multiple 1E records for a person.

Field No	Field Name	Comments
1	Record Type	"1E" - indicates that the record contains telephone numbers or email addresses for the person.
2	Person ID	ID which identifies the person. This ID field must be the same as the ID field in the associated 1A record for the person.
3	Person Phone or email Type	Type of code entered in the following field. Enter the appropriate type code from the table below. The same type may be used more than once. For example, the person may have multiple business phone numbers.
4	Phone Number, Email , or URL	A telephone number, email address, or URL for the person
5	Primary Flag	"Y" if this is the primary phone number or email address for this phone/email type, or the only phone number or email address of this type.. For example, if person phone or email type is "HOME" and the flag is "Y", then this is the primary home phone number.
6	Invalid Flag	"Y" if this person's phone/email address has been marked as invalid/undeliverable.

Person Phone or Email Type Values

Additional types may be added as needed, but the vendor must document all non-standard types that have been used.

Value	Meaning
CELL	Cell phone
EMAIL	Email address
FAX	Fax number
HOME	Home phone number
PAGER	Pager number
PHONE	Phone number, cannot determine the type.
URL	Address of the person's web page
WORK	Work or business phone number

1F Record – Person Attachments

There may be multiple 1F records for a person.

Field No	Field Name	Comments
1	Record Type	"1F" - indicates that the record contains information about documents/files attached to the person record.
2	Person ID	ID which identifies the person. This ID field must be the same as the ID field in the associated 1A record for the person.
3	Document Name	Name of the document/file attached to the person record (for documents/files **not** stored in the form letter/document library). The name must include file path information to identify the location of the document/file.
4	Document ID	ID that identifies the document/file attached to the person record (for documents/files stored in the form letter/document library). This ID must be the same as the Document ID field in the associated 6A record.
5	User ID	User ID of the staffer who attached the document/file. There must be an 8A record containing this staff user ID.
6	Attached Date	Date when the document/file was attached to the person record. Must be in YYYYMMDD format
7	Text	Freeform text containing comments or notes about the attached document/file.

Communication Data

Communication records will be used to identify any way that a person contacts the office, and vice versa (including communications associated with workflows). For example, they might represent a phone call from a constituent, or an outgoing mailing that the office sends out.

Note: Workflow "action history" data is to be provided via 3D – Workflow Text (Comments or Notes)/Action Histories.

2A Record – Communication Data

There may be multiple 2A record for a person.

Field No	Field Name	Comments
1	Record Type	"2A" - Indicates that the record contains communication information.
2	Person ID	ID which identifies the person associated with the communication record. This ID must be the same as the Person ID in the associated 1A record for the person.

Field No	Field Name	Comments
3	Communication ID	A unique ID that identifies the communication.
4	Workflow ID	ID that identifies the workflow associated with this communication. This ID must be the same as the Workflow ID field in the associated 3A record. NULL if the communication is not associated with a workflow.
5	Workflow Person ID	ID that identifies the person associated with the workflow identified in field 4/Workflow ID. This ID must be the same as the Person ID field in the associated 3A record. NULL if the communication is not associated with a workflow.
6	Communication Type	Identifies the type of communication or contact. For example, "LETTER", "EMAIL", "PHONE", "NEWSLETTER", "POST CARD", "VISIT", "DMAIL" (digital mail), etc.
7	User ID	User ID of the individual who handled the communication. There must be an 8A record containing this staff user ID.
8	Approved By	User ID of the staffer who approved this communication. There must be an 8A record containing this staff user ID. NULL if approval is not required, or if approval has not yet been granted.
9	Status*	Status of the communication (e.g. Pending, Completed, Request Approval). *All possible status values must be provided in the table below.
10	Date In	Date communication was received. Must be in YYYYMMDD format.
11	Date Out	Date communication was completed. Must be in YYYYMMDD format.
12	Reminder date	Reminder date for communication. Must be in YYYYMMDD format.
13	Update Date	Date the communication was last updated. Must be in YYYYMMDD format.
14	Response Type	Method of reply. For example, "LETTER", "EMAIL". etc.
15	Address ID	ID which identifies the address to which the communication was sent. This ID field must be the same as the address ID field in the associated 1B record.
16	Email Address	Email address to which the communication was sent.
17	Household Flag	"Y" if the communication was sent to a household. In this case there must also be a 5B record to define the members of the household.
18	Household ID	The household ID in the 5A record which identifies the household. There must be a 5A record containing the household ID.
19	Group Name	Name of the group or batch this communication is a part of.
20	Salutation	The salutation used in the communication.

*Communication Status Values

All possible Communication Status values must be documented in the table below as part of the outgoing vendor's conversion documentation.

Status Value	Closed Flag
(The status provided must be the value that is displayed to the user when viewing the record.)	("Y" if this status indicates that the office has finished/completed their work on this communication, NULL if the status indicates that work on this communication is not completed.)

2B Record – Communication Codes

There may be multiple 2B records for the communication, one for each communication code.

Field No	Field Name	Comments
1	Record Type	"2B" - indicates that the record contains a code associated with the communication.
2	Person ID	ID which identifies the person associated with the communication record. This ID must be the same as the Person ID in the associated 1A record for the person.
3	Communication ID	ID that identifies the communication. This ID must be the same as the Communication ID field in the associated 2A record.
4	Communication Code	A subject or interest code to describe the communication. It can describe the subject of the communication (such as "TAXES") or the constituent's position on a subject (such as "GUNCTRL"). . There must be an 8A record containing this communication subject code.
5	Position	A code that reflects the constituent's position on the subject. For example, "PRO", "CON", "NEUTRAL", or "NONE".

2C Record – Communication Documents

There may be multiple 2C records for the communication, one for each document/file associated with the communication.

Field No	Field Name	Comments
1	Record Type	"2C" - indicates that the record contains document/file name data for the communication.
2	Person ID	ID which identifies the person associated with the communication record. This ID must be the same as the Person ID in the associated 1A record for the person.
3	Communication ID	ID that identifies the communication. This ID must be the same as the Communication ID field in the associated 2A record.
4	Document Type	Identifies whether the following document/file is the incoming document/file, outgoing document/file, or an attachment. See document type values in the table below.
5	Communication Document Name	Name of the communication document/file. The name must include path information to identify the location of the file. For example, "/docs/newsltrs/oct99nl.doc". NULL if the document/file is in the Form Letter/Document Library, in which case the Document ID must be provided in field 6/Communication Document ID.
6	Communication Document ID	ID that identifies the document/file (for example, the ID of the form letter library item sent to the person). This ID must be the same as the Document ID in the associated 6A record. NULL if the document is not in the Form Letter/Document Library, in which case the document/file name/path must be provided in field 5/Communication Document Name.
7	File Location	Location of the document hardcopy (i.e., file cabinet location of the original letter).

Document Type Values

Value	Meaning
INCOMING	Incoming document (i.e. incoming communication from the person)
OUTGOING	Outgoing document (i.e. outgoing communication to the person)
AT_INn	Numbered attachment to the incoming communication, where "n" is the sequence number used to identify the order of the attachments. For example, "AT_IN1" identifies the first incoming attachment, "AT_IN2" identifies the second incoming attachment, etc.
AT_OUTn	Numbered attachment to the outgoing communication, where "n" is the sequence number used to identify the order of the attachments. For example, "AT_OUT1" identifies the first outgoing attachment, "AT_OUT2" identifies the second outgoing attachment, etc.

2D Record – Communication Text (Comments or Notes)

There may be multiple 2D records for the communication.

Field No	Field Name	Comments
1	Record Type	"2D" - indicates comments or notes about the communication.
2	Person ID	ID which identifies the person associated with the communication record. This ID must be the same as the Person ID in the associated 1A record for the person.
3	Communication ID	ID that identifies the communication. This ID must be the same as the Communication ID field in the associated 2A record.
4	2D Sequence Number	Sequence number that identifies the order of 2D records for a communication. A sequence number must be provided if fields 6/Date and 7/Time are NULL.
5	Text	Freeform text containing comments or notes about the communication.
6	Date	Date the text was added to the record. Must be in YYYYMMDD format.
7	Time	Time the text was added to the record. Must be in HH:MM:SS format (e.g., 2:20 PM will be provided as 14:20:00).
8	User ID	User ID of the staffer who entered the text. There must be an 8A record containing this staff user ID.

2E Record – Communication Fill-in Data

When a letter is sent to a person, particularly a form letter, user-defined fill-in fields (variable text inserts) can be included in the letter. The fill-in data may be stored in the letter itself or in a 2E record associated with the outgoing communication.

There may be multiple 2E records for the communication, one for each fill-in field.

Field No	Field Name	Comments
1	Record Type	"2E" - indicates that the record contains a user-supplied fill-in data associated with the communication.
2	Person ID	ID which identifies the person associated with the communication record. This ID must be the same as the Person ID in the associated 1A record for the person.
3	Communication ID	ID that identifies the communication. This ID must be the same as the Communication ID field in the associated 2A record.
4	Fill-in Field Name	The name of the fill-in field as it appears in the document.
5	Fill-in Data	Freeform text that was inserted into the document in place of the fill-in field.

Workflow Data

Data related to an office's business processes, and not defined elsewhere in this format, will be provided as "workflow". Examples of workflow processes are casework (requests for assistance from a constituent regarding a problem/request/inquiry with a federal agency), flag requests, and tour requests.

3A Record – Workflow Data

There may be more than one 3A record for a person if there are multiple workflows for a person.

Field No	Field Name	Comments
1	Record Type	"3A"- Basic workflow record.
2	Person ID	ID which identifies the person about whom the workflow was opened. This ID must be the same as the Person ID in the associated 1A record for the person. NULL if no person is associated with the workflow.
3	Workflow ID	A unique ID that identifies the workflow.
4	Workflow Type	A code that identifies the type of workflow process in this record (e.g. casework, flag request, etc.). Use the appropriate workflow type code from the table below.
5	User ID	User ID of the individual who handled the workflow. There must be an 8A record containing this staff user ID.
6	Start Date	Date workflow was initiated. Must be in YYYYMMDD format.
7	End Date	Date workflow was completed. Must be in YYYYMMDD format.
8	Reminder date	Reminder date for workflow. Must be in YYYYMMDD format.
9	Update Date	Date the workflow was last updated. Must be in YYYYMMDD format.
10	Workflow Description	Freeform description of the workflow. For example, "Veteran benefits inquiry" or "Immigration problem".
11	Status*	The status of the workflow. For example, "Open", "Pending", "Closed Favorably", "Closed Unfavorably". *All possible status values must be provided in the table below.

Workflow Type Values

Additional "workflow types" should be added as needed. All additions must be documented by the vendor.

Value	Meaning
CASE	Casework
PROJECT	Project
GRANT	Grant
FLAG	Flag Request
TOUR	Tour Request
INVITE	Invitation/Meeting Request
APPROP	Appropriation Request
ACADEMY	Academy Nominations
INTERN	Intern Application

*Workflow Status Values

All possible Workflow Status values must be documented in the table below as part of the outgoing vendor's conversion documentation.

Status Value (The status provided must be the value that is displayed to the user when viewing the record.)	Closed Flag ("Y" if this status indicates that the office has finished/completed their work on this workflow, NULL if the status indicates that work on this workflow is not completed.)

3B Record – Workflow Codes

There may be more than one 3B record for a workflow.

Field No	Field Name	Comments
1	Record Type	"3B" - indicates that the record contains codes associated with the workflow.
2	Person ID	ID which identifies the person about whom the workflow was opened. This ID must be the same as the Person ID in the associated 1A record for the person. NULL if no person is associated with the workflow.
3	Workflow ID	ID that identifies the workflow. This ID must be the same as the Workflow ID field in the associated 3A record.

Field No	Field Name	Comments
4	Workflow Code	Code to describe the workflow such as "VET", "SOCSEC", "RAILROAD RETIREMENT" or "IMMIGRATION". There must be an 8A record containing this workflow code.

3D Record – Workflow Text (Comments or Notes)/Action Histories

There may be more than one 3D record for a workflow.

Field No	Field Name	Comments
1	Record Type	"3D" - indicates that the record contains comments or notes associated with the workflow, or an "action history" associated with the workflow.
2	Person ID	ID which identifies the person about whom the workflow was opened. This ID must be the same as the Person ID in the associated 1A record for the person. NULL if no person is associated with the workflow.
3	Workflow ID	ID that identifies the workflow. This ID must be the same as the Workflow ID field in the associated 3A record.
4	3D Sequence Number	Sequence number that identifies the order of 3D records for a workflow. A sequence number must be provided if fields 6/Date and 7/Time are NULL.
5	Text	Freeform text containing comments or notes, or action history details about the workflow.
6	Date	Date the text was added to the record. Must be in YYYYMMDD format.
7	Time	Time the text was added to the record. Must be in HH:MM:SS format (e.g., 2:20 PM will be provided as 14:20:00).
8	User ID	User ID of the staffer who entered the text/action history. There must be an 8A record containing this staff user ID.

3E Record – Workflow Parties

There may be more than one 3E record for a workflow.

Field No	Field Name	Comments
1	Record Type	"3E" - indicates that the record contains information about the persons (e.g. agencies, constituents) associated with the workflow.
2	Person ID	ID which identifies the person about whom the workflow was opened. This ID must be the same as the Person ID in the associated

Field No	Field Name	Comments
		1A record for the person. NULL if no person is associated with the workflow.
3	Workflow ID	ID that identifies the workflow. This ID must be the same as the Workflow ID field in the associated 3A record.
4	Party ID	Person ID of a party (to include agencies) who has an interest in the workflow. There must be a 1A record containing this Person ID.

3F Record – Workflow Attachments

There may be multiple 3F records for a workflow.

Field No	Field Name	Comments
1	Record Type	"3F" - indicates that the record contains information about documents/files attached to the workflow record.
2	Person ID	ID which identifies the person about whom the workflow was opened. This ID must be the same as the Person ID in the associated 1A record for the person. NULL if no person is associated with the workflow.
3	Workflow ID	ID which identifies the workflow. This ID field must be the same as the ID field in the associated 3A record for the workflow.
4	Document Name	Name of the document/file attached to the workflow record (for documents/files **not** stored in the form letter/document library). The name must include file path information to identify the location of the document/file.
5	Document ID	ID that identifies the document/file attached to the workflow record (for documents/files stored in the form letter/document library). This ID must be the same as the Document ID field in the associated 6A record.
6	User ID	User ID of the staffer who attached the document/file. There must be an 8A record containing this staff user ID.
7	Attached Date	Date when the document/file was attached to the workflow record. Must be in YYYYMMDD format
8	Text	Freeform text containing comments or notes about the attached document/file.

3L Record – Additional Workflow Data

The 3L record will be used as a continuation of the 3A record type to transmit workflow type specific field data (fields that are part of the CSS's out-of-the-box configuration) and/or

custom/user-defined field data that is associated with particular workflow types. Further details are provided in the "Additional Requirements" section of this format.

There may be more than one 3L record for a workflow.

Field No	Field Name	Comments
1	Record Type	"3L" - indicates that the record contains additional data about the workflow.
2	Person ID	ID which identifies the person about whom the workflow was opened. This ID must be the same as the Person ID in the associated 1A record for the person. NULL if no person is associated with the workflow.
3	Workflow ID	ID that identifies the workflow. This ID must be the same as the Workflow ID field in the associated 3A record.
4	Workflow Type	A code that identifies the type of workflow process in this record (e.g. casework, flag request, etc.). This Workflow Type must be the same as the Workflow Type provided in the associated 3A record for this workflow.
5	Field Name	Name of the field.
6	Field Value	Value of the field.

Household Data

5A Record – Household Name Data

Field No	Field Name	Comments
1	Record Type	"5A" - indicates that the record contains household name information.
2	Household ID	A unique ID which identifies the household.
3	Household Name	Name of the household (e.g., "The Smith Family")
4	Household Salutation	The salutation used for the household (e.g., "Smith Family").

5B Record – Household Member Data

There will be multiple 5B records for a household, one for each person/member of the household.

Field No	Field Name	Comments
1	Record Type	"5B" - indicates that the record contains household member information.
2	Household ID	ID which identifies the household. This ID must be the same as the ID field in the associated 5A record for the household.
3	Person ID	A Person ID that identifies a person who is a member of the household. There must be a 1A record containing this Person ID.
4	Primary Contact Flag	"Y" if this person is the head of household or primary contact.

5C Record – Household Attachments

There may be multiple 5C records for a Household.

Field No	Field Name	Comments
1	Record Type	"5C" - indicates that the record contains information about documents/files attached to a household record.
2	Household ID	ID which identifies the household. This ID must be the same as the ID field in the associated 5A record for the household.
3	Document Name	Name of the document/file attached to the household record (for documents/files **not** stored in the form letter/document library). The name must include file path information to identify the location of the document/file.
4	Document ID	ID that identifies the document/file attached to the household record (for documents/files stored in the form letter/document library). This ID must be the same as the Document ID field in the associated 6A record.
5	User ID	User ID of the staffer who attached the document/file. There must be an 8A record containing this user ID.
6	Attached Date	Date when the document/file was attached to the household. Must be in YYYYMMDD format
7	Text	Freeform text containing comments or notes about the attached document/file.

5D Record – Household Text (Comments or Notes)

The may be multiple 5D records for a Household.

Field No	Field Name	Comments
1	Record Type	"5D" - indicates that the record contains comments or notes associated with the household.
2	Household ID	ID which identifies the household. This ID must be the same as the ID field in the associated 5A record for the household.
3	5D Sequence Number	Sequence number that identifies the order of 5D records for a household. A sequence number must be provided if fields 5/Date and 6/Time are NULL.
4	Text	Freeform text containing comments or notes about the household.
5	Date	Date the text was added to the record. Must be in YYYYMMDD format.
6	Time	Time the text was added to the record. Must be in HH:MM:SS format (e.g., 2:20 PM will be provided as 14:20:00).
7	User ID	User ID of the staffer who entered the text. There must be an 8A record containing this Staff/user ID.

Form Letter/Document Library Data
6A Record – Form Letter/Document Library Data

The following table will be used to define form letter library items and documents/files that are part of an office's document library. Files in the document library can be any document/file (most likely with meta-data attached) that an office wishes to retain for historical or reference purposes. As such, incoming email messages, incoming correspondence images, ad-hoc outgoing correspondence, and the like are not considered to be part of the document library.

Field No	Field Name	Comments
1	Record Type	"6A" – indicates that the record contains form letter/document data.
2	Document ID	A unique ID that identifies this version of the form letter/document.
3	Version	Version of the form letter/document. (e.g. 1, 2, 3)
4	Document Grouping ID	ID which links multiple versions of a form letter/document.
5	Document Type	The type of document/file described in this record. "FORM" for form letter library item, or "DOC" for document library file.
6	Document Display Name	Display name of this version of the form letter/document (i.e. the name that the user sees). Must be provided for all 6A records.
7	Document Description	Freeform text describing the form letter/document.
8	Document Name	Name of this version of the form letter/document file. The name must include path/directory structure information to identify the physical location of the file. For example, "/docs/newsltrs/oct99nl.doc".

Field No	Field Name	Comments
9	Created By	User ID of the staffer who created this version of the form letter/document. There must be an 8A record containing this staff user ID.
10	Revised By	User ID of the staffer who last revised this version of the form letter/document. There must be an 8A record containing this staff user ID.
11	Approved By	User ID of the staffer who approved this version of the form letter/document. There must be an 8A record containing this staff user ID. NULL if approval is not required, or if approval has not yet been granted.
12	Creation Date	Creation date for this version of the form letter/document. Must be in YYYYMMDD format.
13	Revision Date	Last revision date for this version of the form letter/document. Must be in YYYYMMDD format.
14	Last Used Date	Date that this version of the form letter/document was last used/assigned. Must be in YYYYMMDD format.
15	Status	Status of this version of the form letter/document (e.g. Draft, Approved). .
16	Inactive Flag	"Y" if this version of the form letter/document is inactive (for example, a form letter library item that has been used/assigned previously, but is not currently available for use because it's content is out of date)
17	Virtual Directory	Virtual directory structure/path of the form letter/document (i.e. the directory structure/path that the user sees). For example, "\FY16 Appropriation Requests\Education\".

6B Record – Form Letter/Document Library Fill-in Data

There may be multiple 6B records for a form letter/document, one for each fill-in (variable text insert) field.

Field No	Field Name	Comments
1	Record Type	"6B" – indicates that the record contains form letter/document fill-in data.
2	Document ID	ID that identifies this version of the form letter/document. This ID must be the same as the Document ID field in the associated 6A record.
3	Fill-in Field Name	The name of the fill-in field as it appears in the document/file.

6C Record – Form Letter/Document Library Code Data

This record is used to associate a form letter/document with workflow, communication or document codes. There may be multiple 6C records for a form letter/document.

Field No	Field Name	Comments
1	Record Type	"6C" – indicates that the record contains form letter/document code data.
2	Document ID	ID that identifies this version of the form letter/document. This ID must be the same as the Document ID field in the associated 6A record.
3	Code Type	Type of code entered in the following field. Enter the appropriate code type from the table below.
4	Code	Code used to describe the form letter/document. There must be an 8A record containing this code.

Code Type Values

Value	Meaning
WORK	Workflow codes
COM	Communication subject codes
DOC	Document codes

6D Record – Form Letter/Document Library Attachments

There may be multiple 6D records for a Form Letter/Document.

Field No	Field Name	Comments
1	Record Type	"6D" - indicates that the record contains information about documents/files attached to a form letter/document record.
2	Document ID	ID that identifies this version of the form letter/document. This ID must be the same as the Document ID field in the associated 6A record.
3	Document Name	Name of the document/file attached to the form letter/document. The name must include file path information to identify the location of the document/file.
4	User ID	User ID of the staffer who attached the document/file. There must be an 8A record containing this user ID.
5	Attached Date	Date when the document/file was attached to the form letter/document. Must be in YYYYMMDD format
6	Text	Freeform text containing comments or notes about the attached document/file.

Field No	Field Name	Comments
7	Form Letter Attachment Flag	"Y" if this document/file is attached to a form letter **and** should be sent as part of the form letter response.

6E Record – Form Letter/Document Library Text (Comments or Notes)

The may be multiple 6E records for a Form Letter/Document.

Field No	Field Name	Comments
1	Record Type	"6E" - indicates that the record contains comments or notes associated with the form letter/document.
2	Document ID	ID that identifies this version of the form letter/document. This ID must be the same as the Document ID field in the associated 6A record.
3	6E Sequence Number	Sequence number that identifies the order of 6E records for a form letter/document. A sequence number must be provided if fields 5/Date and 6/Time are NULL.
4	Text	Freeform text containing comments or notes about the form letter/document.
5	Date	Date the text was added to the record. Must be in YYYYMMDD format.
6	Time	Time the text was added to the record. Must be in HH:MM:SS format (e.g., 2:20 PM will be provided as 14:20:00).
7	User ID	User ID of the staffer who entered the text. There must be an 8A record containing this Staff/user ID.

6F Record – Form Letter/Document Library Owner Data

There may be multiple 6F records for a form letter/document.

Field No	Field Name	Comments
1	Record Type	"6F" – indicates that the record contains form letter/document owner data.
2	Document ID	ID that identifies this version of the form letter/document. This ID must be the same as the Document ID field in the associated 6A record.
3	Owned By	User ID of the owner of this version of the form letter/document (may also be known as the responsible staffer or the staffer that this form letter/document is assigned to). There must be an 8A record containing this staff user ID.

Schedule Data
7A Record – Schedule/Event Data

Field No	Field Name	Comments
1	Record Type	"7A" - indicates that the record contains schedule data.
2	Event ID	A unique ID which identifies the event.
3	Recurrence ID	ID which identifies the recurrence if this event is part of a recurring series. There must be a 7H record containing this recurrence ID. NULL if this event is not part of a recurring series. (Note: each event that is part of a recurring series must be provided individually in the 7A-F records)
4	Event Description	Freeform text describing the event.
5	Contact	Contact information about the person requesting/organizing the event (name, phone number, etc.)
6	County	County in which the event takes place
7	Organization	Organization hosting the event
8	Start Date	Event start date in YYYYMMDD format.
9	Start Time	Event start time in HH:MM:SS format (e.g., 2:20 PM will be provided as 14:20:00). The time provided will be the time in the time zone in which the event occurs.
10	End Date	Event end date in YYYYMMDD format.
11	End Time	Event end time in HH:MM:SS format. The time provided will be the time in the time zone in which the event occurs.
12	Time Zone	Time Zone in which the event takes place (EST, CST, MST, PST, etc). The default is EST.
13	Location	The location of the event (e.g. the DC or state office).
14	Status	Status (e.g. Pending, Approved, Tentative, Declined)
15	Private Flag	"Y" if event is marked as private.
16	Scheduled Date	Date the event was entered into the schedule. Must be in YYYYMMDD format.
17	Scheduled by	User ID of staff person entering the event. There must be an 8A record containing this user ID.
18	Revision Date	Date the event was last revised. Must be in YYYYMMDD format.
19	Revised By	User ID of the staff person who last revised the event. There must be an 8A record containing this staff user ID.
20	Address line 1	First line of address at which the event takes place.
21	Address line 2	Second line of address at which the event takes place.
22	Address line 3	Third line of address at which the event takes place.
23	Address line 4	Fourth line of address at which the event takes place.
24	City	City in which the event takes place.
25	State Code	2 letter state code for the state in which the event takes place. For example, "VA", "DC" or "MD".

Field No	Field Name	Comments
26	Zip Code	Zip code in which the event takes place. Format for 10 digit zip code is XXXXX-XXXX. If only the 5-digit code is known, then the format is XXXXX. If international address, enter the appropriate zip code of the country.
27	Country	Country (if international address) in which the event takes place.

7B Record – Event Code Data

There may be more than one 7B record for an event.

Field No	Field Name	Comments
1	Record Type	"7B" – indicates that the record contains schedule/event code data.
2	Event ID	ID that identifies the event. This ID must be the same as the Event ID field in the associated 7A record.
3	Event Code	A subject or interest code to describe the event. There must be an 8A record containing this event code.

7C Record – Event Participant Data

There may be more than one 7C record for an event.

Field No	Field Name	Comments
1	Record Type	"7C" – indicates that the record contains schedule event participant data.
2	Event ID	ID that identifies the event. This ID must be the same as the Event ID field in the associated 7A record.
3	Participant Type	Type of event participant entered in the following field. Enter the appropriate participant type code from the table below.
4	Participant ID	User ID of the Member or staffer who is to participate in the event. There must be an 8A record containing this user ID. OR Person ID of the person who is to participate in the event. There must be a 1A record containing this Person ID.

Participant Type Values

Value	Meaning
STAFF	Member or staff who is to participate in the event.
PERSON	Person who is to participate in the event.

7D Record – Event Attachments

There may be multiple 7D records for an event.

Field No	Field Name	Comments
1	Record Type	"7D" - indicates that the record contains information about documents/files attached to a schedule event record.
2	Event ID	ID which identifies the event. This ID field must be the same as the ID field in the associated 7A record for the event.
3	Document Name	Name of the document/file attached to the event (for documents/files **not** stored in the form letter/document library). The name must include file path information to identify the location of the document/file.
4	Document ID	ID that identifies the document/file attached to the event (for documents/files stored in the form letter/document library). This ID must be the same as the Document ID field in the associated 6A record.
5	User ID	User ID of the staffer who attached the document/file. There must be an 8A record containing this staff user ID.
6	Attached Date	Date when the document/file was attached to the event. Must be in YYYYMMDD format
7	Text	Freeform text containing comments or notes about the attached document/file.

7E Record – Event Text (Comments or Notes)

The may be multiple 7E records for an Event.

Field No	Field Name	Comments
1	Record Type	"7E" - indicates that the record contains comments or notes associated with the event.
2	Event ID	ID which identifies the event. This ID field must be the same as the ID field in the associated 7A record for the event.
3	7E Sequence Number	Sequence number that identifies the order of 7E records for an event. A sequence number must be provided if fields 5/Date and 6/Time are NULL.
4	Text	Freeform text containing comments or notes about the event.
5	Date	Date the text was added to the record. Must be in YYYYMMDD format.
6	Time	Time the text was added to the record. Must be in HH:MM:SS format (e.g., 2:20 PM will be provided as 14:20:00).

Field No	Field Name	Comments
7	User ID	User ID of the staffer who entered the text. There must be an 8A record containing this Staff/user ID.

7F Record – Event Associated Workflows

There may be multiple 7F records for an event.

Field No	Field Name	Comments
1	Record Type	"7F" - indicates that the record contains information about workflows associated with a schedule event.
2	Event ID	ID which identifies the event. This ID field must be the same as the ID field in the associated 7A record for the event.
3	Workflow ID	ID that identifies the workflow attached to the event. This ID must be the same as the Workflow ID field in the associated 3A record.

7H Record – Recurring Series Definition Data

Field No	Field Name	Comments
1	Record Type	"7H" – indicates that the record contains recurring series definition data.
2	Recurrence ID	A unique ID that identifies the recurring series.
3	Start Series Date	Date of the first event for the series of recurring events. Must be in YYYYMMDD format.
4	End Series Date	Date of the last event for the series of recurring events. Must be in YYYYMMDD format. (NULL, if no "end series date" is specified.)
5	Recurrence Type	"DAILY", "WEEKLY", "MONTHLY" or "YEARLY"
6	Recurrence Interval	Definition of the recurrence interval. Values are per the recurrence interval schema*** below.

Recurrence Interval Schema***

Recurrence Type	Recurrence Interval Definition
DAILY	Pattern of the "daily" recurrence. Defined with one value. Defines the daily frequency interval. Valid values: "1", "2", ..., "weekdays", "weekenddays".
WEEKLY	Pattern of the "weekly" recurrence. Defined with two values separated by a dash. Part one defines the weekly frequency interval. Valid values: "1", "2," …. Part two defines the day of the week. Valid values: "sunday", "monday", "tuesday", "wednesday", "thursday", "friday", "saturday"
MONTHLY	Pattern of the "monthly" recurrence. Defined with three values separated by dashes. Part one defines the monthly frequency interval. Valid values: "1", "2," …. Part two defines the day descriptor. Valid values: "1", "2", …, "31", "last". Part three defines the type of day. Valid values: "sunday", "monday", "tuesday", "wednesday", "thursday", "friday", "saturday", "day", "weekday", "weekendday".
YEARLY	Pattern of the "yearly" recurrence. Defined with four values separated by dashes. Part one defines the yearly frequency interval. Valid values: "1", "2," …. Part two defines the day descriptor. Valid values: "1", "2", …, "31", "last". Part three defines the type of day. Valid values: "sunday", "monday", "tuesday", "wednesday", "thursday", "friday", "saturday", "day", "weekday", "weekendday". Part four defines the month. Valid values: "january", "february", "march", "april", "may", "june", "july", "august", "september", "october", "november", "december".

Sample Recurrence Interval Values

Value	Meaning
1	Event recurs every day (for "DAILY" recurrence type).
4	Event recurs every 4th day (for "DAILY" recurrence type)
weekdays	Event recurs every weekday (for "DAILY" recurrence type).
weekenddays	Event recurs every weekend day (for "DAILY" recurrence type).
1-sunday	Event recurs every week on Sunday (for "WEEKLY" recurrence type).
2-monday	Event recurs every 2nd week on Monday (for "WEEKLY" recurrence type).
1-1-day	Event recurs every month on the 1st day of the month (for "MONTHLY" recurrence type).
3-15-day	Event recurs every 3rd month on the 15th day of the month (for "MONTHLY" recurrence type).
2-last-day	Event recurs every 2nd month on the last day of the month (for "MONTHLY" recurrence type).
4-1-weekday	Event recurs every 4th month on the first weekday of the month (for "MONTHLY" recurrence type).
3-1-weekendday	Event recurs every 3rd month on the first weekend day of the month (for "MONTHLY" recurrence type
1-1-tuesday	Event recurs every month on the 1st Tuesday of the month (for "MONTHLY" recurrence type).
4-3-sunday	Event recurs every 4th month on the 3rd Sunday of the month (for "MONTHLY" recurrence type).
1-9-day-may	Event recurs every year on the 9th day of May (for the "YEARLY" recurrence type).
1-5-weekday-june	Event recurs every year on the 5th weekday of June (for the "YEARLY" recurrence type).
1-last-day-january	Event recurs every year on the last day of January (for the "YEARLY" recurrence type).
2-1-tuesday-november	Event recurs every 2nd year on the 1st Tuesday of November (for the "YEARLY" recurrence type).
1-2-weekendday-april	Event recurs every year on the 2nd weekend day of April (for the "YEARLY" recurrence type).

The 7I – 7N records will be used as adjuncts to the 7H record (Recurring Series Definition Data) and will be used to transmit any additional data associated with a recurring series.

7I Record – Recurring Series Source Data

Field No	Field Name	Comments
1	Record Type	"7I" - indicates that the record contains recurring series source data.
2	Recurrence ID	A unique ID that identifies the recurring series. This ID must be the same as the Recurrence ID field in the associated 7H record.
3	Not used	Not used
4	Event Description	Freeform text describing the recurring series.
5	Contact	Contact information about the person requesting/organizing the recurring series (name, phone number, etc.)
6	County	County in which the recurring series takes place
7	Organization	Organization hosting the recurring series
8	Not used	Not used
9	Start Time	Recurring series start time in HH:MM:SS format (e.g., 2:20 PM will be provided as 14:20:00). The time provided will be the time in the time zone in which the recurring series occurs.
10	Not used	Not used.
11	End Time	Recurring series end time in HH:MM:SS format. The time provided will be the time in the time zone in which the recurring series occurs.
12	Time Zone	Time Zone in which the recurring series takes place (EST, CST, MST, PST, etc). The default is EST.
13	Location	The location of the recurring series (e.g. the DC or state office).
14	Status	Status (e.g. Pending, Approved, Tentative, Declined)
15	Private Flag	"Y" if recurring series is marked as private.
16	Scheduled Date	Date the recurring series was entered into the schedule. Must be in YYYYMMDD format.
17	Scheduled by	User ID of staff person entering the recurring series. There must be an 8A record containing this user ID.
18	Revision Date	Date the recurring series was last revised. Must be in YYYYMMDD format.
19	Revised By	User ID of the staff person who last revised the recurring series. There must be an 8A record containing this staff user ID.
20	Address line 1	First line of address at which the recurring series takes place.
21	Address line 2	Second line of address at which the recurring series takes place.
22	Address line 3	Third line of address at which the recurring series takes place.
23	Address line 4	Fourth line of address at which the recurring series takes place.
24	City	City in which the recurring series takes place.
25	State Code	2 letter state code for the state in which the recurring series takes place. For example, "VA", "DC" or "MD".
26	Zip Code	Zip code in which the recurring series takes place. Format for 10 digit zip code is XXXXX-XXXX. If only the 5-digit code is known, then the format is XXXXX. If international address, enter the appropriate zip code of the country.

Field No	Field Name	Comments
27	Country	Country (if international address) in which the recurring series takes place.

7J Record – Recurring Series Code Data

There may be more than one 7J record for a recurring series.

Field No	Field Name	Comments
1	Record Type	"7J" – indicates that the record contains recurring series code data.
2	Recurrence ID	A unique ID that identifies the recurring series. This ID must be the same as the Recurrence ID field in the associated 7H record.
3	Event Code	A subject or interest code to describe the recurring series. There must be an 8A record containing this event code.

7K Record – Recurring Series Participant Data

There may be more than one 7K record for a recurring series.

Field No	Field Name	Comments
1	Record Type	"7K" – indicates that the record contains recurring series participant data.
2	Recurrence ID	A unique ID that identifies the recurring series. This ID must be the same as the Recurrence ID field in the associated 7H record.
3	Participant Type	Type of recurring series participant entered in the following field. Enter the appropriate participant type code from the table below.
4	Participant ID	User ID of the Member or staffer who is to participate in the recurring series. There must be an 8A record containing this user ID. OR Person ID of the person who is to participate in the recurring series. There must be a 1A record containing this Person ID.

Participant Type Values

Value	Meaning
STAFF	Member or staff who is to participate in the recurring series.
PERSON	Person who is to participate in the recurring series.

7L Record – Recurring Series Attachments

There may be multiple 7L records for a recurring series.

Field No	Field Name	Comments
1	Record Type	"7L" - indicates that the record contains information about documents/files attached to a recurring series record.
2	Recurrence ID	A unique ID that identifies the recurring series. This ID must be the same as the Recurrence ID field in the associated 7H record.
3	Document Name	Name of the document/file attached to the recurring series (for documents/files **not** stored in the form letter/document library). The name must include file path information to identify the location of the document/file.
4	Document ID	ID that identifies the document/file attached to the recurring series (for documents/files stored in the form letter/document library). This ID must be the same as the Document ID field in the associated 6A record.
5	User ID	User ID of the staffer who attached the document/file. There must be an 8A record containing this staff user ID.
6	Attached Date	Date when the document/file was attached to the recurring series. Must be in YYYYMMDD format
7	Text	Freeform text containing comments or notes about the attached document/file.

7M Record – Recurring Series Text (Comments or Notes)

The may be multiple 7M records for a recurring series.

Field No	Field Name	Comments
1	Record Type	"7M" - indicates that the record contains comments or notes associated with the recurring series.
2	Recurrence ID	A unique ID that identifies the recurring series. This ID must be the same as the Recurrence ID field in the associated 7H record.
3	7M Sequence Number	Sequence number that identifies the order of 7M records for a recurring series. A sequence number must be provided if fields 5/Date and 6/Time are NULL.
4	Text	Freeform text containing comments or notes about the recurring series.
5	Date	Date the text was added to the record. Must be in YYYYMMDD format.
6	Time	Time the text was added to the record. Must be in HH:MM:SS format (e.g., 2:20 PM will be provided as 14:20:00).

Field No	Field Name	Comments
7	User ID	User ID of the staffer who entered the text. There must be an 8A record containing this Staff/user ID.

7N Record – Recurring Series Associated Workflows

There may be multiple 7N records for a recurring series.

Field No	Field Name	Comments
1	Record Type	"7N" - indicates that the record contains information about workflows associated with a recurring series.
2	Recurrence ID	A unique ID that identifies the recurring series. This ID must be the same as the Recurrence ID field in the associated 7H record.
3	Workflow ID	ID that identifies the workflow attached to the recurring series. This ID must be the same as the Workflow ID field in the associated 3A record.

Code Table Data

8A Record – Code Table Data

Field No	Field Name	Comments
1	Record Type	"8A" – indicates that the record contains code table data.
2	Code Type	The type of code entered in the following field. Code types are defined in the table below.
3	Code	A standard code used throughout this format.
4	Code Description	A description or value of the code. If the code description is not present in the source data, the value of the code itself shall be repeated in this field.
5	Inactive Flag	"Y" if the code is inactive (i.e. the code has been assigned previously, but is not currently available for assignment).
6	User ID	Where applicable, the User ID of the staffer associated with this code. There must be an 8A record containing this Staff/user ID.

Code Type Values

Value	Record Location	Type Code
WORK	1C, 3B, 6C	Workflow codes
	1C	Person codes

Value	Record Location	Type Code
PERS		
COM	2B, 6C	Communication subject codes
EVEN T	7B	Event codes
DOC	6C	Document codes
STAFF	1D,1F, 2A, 2D, 3A, 3D, 3F, 5C, 5D, 6A, 6D, 6E, 7A, 7C, 7D, 7E	Staff user IDs

The record location indicates the types of records in which the code is located.

Additional Requirements

Documents/Attachments

All documents/files associated with the CSS will be copied and provided to the recipient. All directory structures, including virtual directories, must be maintained. All documents/files must include file extensions, and the file extensions must accurately reflect the format of the file.

Word processing documents/files (responses to communications, form letter library documents, word processing documents in the document library, etc) will be provided in a format such that all formatting is maintained when the file is opened in Word.

Incoming email message files will be provided in their native format (e.g. .txt, .eml).

All other documents/files will be provided in their native format (e.g. .xls, .tif, .pdf,).

Custom/User-defined Fields

Custom/user defined fields are fields that are unique to a particular office's CSS installation, and are not part of the standard configuration of the CSS. Custom/user-defined fields will be added to the format and documented.

Additional Workflow Data

CSS specific field data that is associated with particular workflow types (i.e. some, but not all, workflow types/templates) will be provided in the "3L – Additional Workflow Data" record. This data will be documented within the format.

Conversion Documentation

Updates/additions to the format, where permitted, will be documented within the format (the new fields must be added to the existing format and highlighted). Documentation will include:
- the field number (not applicable for 3L – Additional Workflow Data),
- the field name,
- a description of the data (the description must expand on the field name and provide enough detail to allow the receiving vendor to understand the purpose of data),
- a notation identifying custom/user-defined fields, as applicable,
- the data type (e.g. date, time, number or text),
- all pre-defined data values, as applicable, and
- for workflow data, the "workflow type code" that identifies the type of workflow process the field is used with.

Sample documentation for addition of a custom field:

1A Record – Name Data

Field No	Field Name	Comments
1	Record Type	"1A" - indicates that the record contains person name information
2	Person ID	A unique numeric ID which identifies the person.
3	Person Type	A code that identifies the type of person in this record. Use the appropriate entry from the table below.
4	Prefix	Prefix of person such as "Mr.", "Mrs." or "The Honorable".
5	First Name	First name of person.
6	Middle Name	Middle name or initial of person.
7	Last Name	Last name of person.
8	Suffix	Suffix of person such as "Jr.".
9	Appellation	Appellation of person such as "MD".
10	Salutation	Preferred salutation used for the person, such as "Bob" or " Senator Jones".
11	Date of Birth	The person's date of birth in the format YYYYMMDD.
12	No Mail Flag	"Y" if the person has requested that no mail be sent to then at any of their addresses.
13	Deceased Flag	"Y" if the person is deceased.
14	Spouse's Name	Name of the person's spouse.
15	Email Flag	"Y" if the person has requested that their preferred method of communication is via email.
16	Gender	Gender of the person. Custom field. Data format=text. Valid values: "MALE" or "FEMALE".

Sample documentation for additional workflow data:

3L Record -- Additional Workflow Data – Workflow Type Code = CASE

Field Name	Description	Custom Field?	Data Type	Possible Values
SSN	The person's social security number.	Yes	Text	

3L Record -- Additional Workflow Data – Workflow Type Code = FLAG

Field Name	Description	Custom Field?	Data Type	Possible Values
Flag Type	Describes the type of flag requested.		Text	3x5 cotton 3x5 nylon 5x8 cotton 5x8 nylon
Fly Date	The date the person has requested the flag to be flown over the Capitol.		Date	
Occasion	Describes the occasion for which the flag request was made (for example, birthday, graduation, etc).		Text	

3L Record -- Additional Workflow Data – Workflow Type Code = TOUR

Field Name	Description	Custom Field?	Data Type	Possible Values
Group Size	The number of people in the group.		Number	
Arrival Date	The date the group arrives in DC.		Date	
Departure Date	The date the group departs from DC.		Date	
Special Needs	Free form text describing any special needs of the group.		Text	
White House Tour Flag	"Y" if a White House tour is requested.		Text	Y NULL
Capitol Tour Flag	"Y" if a Capitol tour is requested.		Text	Y NULL

Sample documentation for "status" values:

*Communication Status Values

Status Value	Closed Flag
(The status provided must be the value that is displayed to the user when viewing the record.)	("Y" if this status indicates that the office has finished/completed their work on this communication, NULL if the status indicates that work on this communication is not completed.)
PENDING	
REQUEST APPROVAL	
CLOSED	Y

*Workflow Status Values

Status Value	Closed Flag
(The status provided must be the value that is displayed to the user when viewing the record.)	("Y" if this status indicates that the office has finished/completed their work on this workflow, NULL if the status indicates that work on this workflow is not completed.)
OPEN	
CLOSED FAVORABLY	Y
CLOSED	Y

"Date" and "Time" Fields

All "date" values will be provided in YYYYMMDD format.

All "time" values will be provided in HH:MM:SS format (e.g. 2:07pm and 23 seconds will be provided as 14:07:23).

"Type Codes/Values"

New type codes/values may be added only in those areas specifically noted within the format (i.e. 1B "address type values", 1E "person phone or email type values", 3A "workflow type values"). All such additions must be fully documented.

APPENDIX D: DETAILED CSS/CMS SURVEY RESULTS

In February 2017, members of the CSS/CMS Task Force reached out to repositories that house data in their congressional collections in an effort to gather qualitative information on work that has been done with CSS/CMS data, as well as any obstacles encountered. The Task Force received responses from three institutions (respondents A, B & C) which are custodians of at least eleven sets of CSS/CMS data.

Within respondent A's repository, accessing and preserving systems data (and digital materials in general) had historically been a low priority. Previous archivists had discouraged congressional offices from transferring systems data in digital form. As such, the repository retains systems data on media used for system backups (hard drives, data reels, and data cassettes) as well as in paper form. Respondent B was not yet employed at their institution at the time of transfer, but Respondent C did indicate some level of communications with the originating offices. They described a delay in receiving the data, stemming from both the vendor export process and the Senate Sergeant at Arms' review process, as well as an unsuccessful attempt to gather a full export of data from one office.

Of the two respondents whose repositories contain systems data in digital form, Respondent B received a single set of data on a DVD burned by the vendor, Lockheed Martin. (The originating office used Intranet Quorum). The repository currently maintains multiple copies of electronic files for preservation purposes, and are working to establish workflows for long-term digital preservation. They were able to import the data into an SQL database, but Respondent B described it as a "huge data dump" containing a significant amount of personally identifiable information (PII). No further work has been performed, but Respondent B is actively engaging in conversations with other repositories to find solutions for preservation and access.

Respondent C reported six sets of systems data, three from Fireside 21 and three whose software origins are unknown. Data was received largely on CDs, though one office sent data on an external hard drive. Most of the metadata arrived in .DAT files with emails in .TXT files, but one set of metadata arrived in .CSV files with emails zipped in .RAR files and two sets are still on original media. Respondent C experimented with importing one set of data into MySQL, adapting a search interface they had made for a class project to work with the database. During the process, however, they realized that the data had initially been exported incorrectly by the vendor, and that the identifier linking the metadata to the emails was missing. Since it was not a complete data set, they did not pursue this further. The respondent's repository currently has secure server space to store files, as well as a preliminary preservation workflow and policies, all of which have guided work with the systems data.

All respondents reported that the systems data in their collections are closed, even those in open collections, due to the large amount of PII and the inability to properly review and redact such large quantities of data. Resource constraints also place severe limitations on working with systems data: Repositories which lack staff, time, frameworks for digital preservation, and the equipment/expertise needed to transfer data from physical media to digital formats are at a serious disadvantage.

Respondents clearly stated their desire for an open-source database API which would facilitate the ingest of data, contain PII identification and redaction capabilities, and provide mediated access to researchers. They also requested visualization capabilities related to correspondence subjects, as well as a tool to allow for the aggregation and comparison of multiple data sets. One respondent expressed a desire for researchers to be able to download an anonymized set of data, as well as the ability to identify and suppress or redact restricted information. More generally, respondents indicated the need for a tool which can be implemented and used with minimal assistance from IT staff, an important consideration for repositories with constrained resources.

Most 2014 survey respondents expressed a desire for more data as a part of a records export from vendors, including a greater number of fields (the survey was distributed prior to the release of the SCDIF format export option); full metadata related to storage and access; and data related to scheduling and legislation. Respondents also expressed interest in receiving contextual information related to data structure, as well as any information related to decision-making processes regarding procedures and systems software/vendors. While it is difficult to draw conclusions from such a small pool of data, these responses highlight trends in the experience of archivists working with CSS/CMS data across the nation.

BIBLIOGRAPHY

Advisory Committee on the Records of Congress, meeting 46 minutes, June 16, 2014. https://www.archives.gov/files/legislative/cla/advisory-committee/minutes/2014-19-06.pdf

Allen, Neal. "Research in Congressional Holdings," presentation to the Association of Centers for the Study of Congress 12th annual meeting, Washington, DC, May 13-15, 2015.

"Bart Gordon Papers." Albert Gore Research Center, Middle Tennessee State University, Murfreesboro, TN, 2013. http://gorecenter.mtsu.edu/research/BartGordon.shtml.

Bell, Joni. "Congressional Demographics – It's [sic] Effect on Member of Congress Files." Society of American Archivists Congressional Papers Roundtable newsletter, December 1987, 5-7. https://www2.archivists.org/sites/all/files/1987_December.pdf.

Bond, Natalie. "Testing the Waters: Working with CSS Data in Congressional Collections." *Society of American Archivists Congressional Papers Roundtable Electronic Records Case Studies Series*. August 2015. https://cprerc.wordpress.com/case-studies/case-study-004/.

"Computers on Capitol Hill: SAA in NY." Society of American Archivists Congressional Papers Roundtable newsletter, December 1987, 4. https://www2.archivists.org/sites/all/files/1987_December.pdf.

"Congressional Papers Roundtable Minutes 1998 Annual Meeting Orlando, September 4." Society of American Archivists Congressional Papers Roundtable newsletter, November 1998, 1-4. https://www2.archivists.org/sites/all/files/1998_November.pdf.

Fried, Amy. "Public Opinion in the Archives." In *Doing Archival Research in Political Science*, edited by Scott A. Frisch, Douglas B. Harris, Sean Q. Kelly, and David C.W. Parker, Amherst, NY: Cambria Press, 2012, and David Thelan, *Becoming Citizens in the Age of Television: How Americans Challenged the Media and Seized Political Initiative During the Iran-Contra Debate*. Chicago: The University of Chicago Press, 1996.

Fry, Joseph A. *The American South and the Vietnam War*. Kentucky: The University of Kentucky Press, 2015.

Hanson, Adriane. "Donations of Digital Records from Congressional Offices: Lessons Learned from the 2014 Election Cycle," *Society of American Archivists Congressional Papers Roundtable Electronic Records Case Studies Series*. June 2015. https://cprerc.wordpress.com/case-studies/donations-of-digital-records-from-congressional-offices-june-2015/.

Nelson, Naomi. "From 'ROBO' Letters to E-mail: The Evolution of the Constituent Mail Function in the Senate." Society of American Archivists Congressional Papers Roundtable newsletter, April 1995, 3-5. https://www2.archivists.org/sites/all/files/1995_April.pdf.

Nelson, Naomi. "Taking a Byte Out of the Senate: Reconsidering the Research Use of Correspondence and Casework Files." In *An American Political Archives Reader*, edited by Karen Dawley Paul, Glenn R. Gray, and L. Rebecca Johnson Melvin, 235-252, Lanham, Maryland: The Scarecrow Press, Inc., 2009.

Perino, Michael. *The Hellhound of Wall Street: How Ferdinand Pecora's Investigation of the Great Crash Forever Changed American Finance.* New York: The Penguin Press, 2010.

Scott, Katherine. "A Safety Valve: The Truman Committee's Oversight during World War II." In *Congress and Civil-Military Relations*, edited by Colton C. Campbell and David P. Auerswald, Washington, D.C.: Georgetown University Press, 2015.

"Rockefeller-CSS." West Virginia University Libraries, 2016. https://github.com/wvulibraries/rockefeller-css.

United States Senate. Office of the Sergeant at Arms. CSS III Solicitation 20160R-025, 2016.

United States Senate. *Records Management Handbook for United States Senators and Their Repositories,* by Karen Dawley Paul, Y 1.3:S.pub.102-17, Washington, D.C.: GPO, 1992.

Zaretsky, Natasha. "Restraint or Retreat? The Debate over the Panama Canal Treaties and U.S. Nationalism after Vietnam." *Diplomatic History*, 35, no. 3 (June 2011).