

Needs Assessment Survey Analysis and Report

Prepared by Nick Connizzo and Michelle Gallinger
State Electronic Records Initiative



RECENT REPORTS AND PUBLICATIONS FROM THE COUNCIL OF STATE ARCHIVISTS

State of State Records: A Statistical Report on State Archives and Records Management Programs in the United States. JULY 2021

SERI Digital Best Practices Series Number 2: Developing Successful Electronic Records Grants Projects. MAY 2021

MoVE-IT: Modeling Viable Electronic Information Transfers. FEBRUARY 2021

The Modeling Viable Electronic Information Transfers (MoVE-IT) project builds upon work surveying state and territorial archives regarding the efficacy of their electronic records management and digital preservation programs. This report analyzes seven electronic records transfer projects with regard to the elements that challenged them and, ultimately, made them successes. Published with support from Preservica and AVP.

Public Records and Remote Work: Advice for State Governments. SEPTEMBER 2020

Social Media as State Public Records: A Collections Practices Scan. MAY 2020

This report clarifies current state archives collection practices of state electronic records on social media platforms, providing a window into how state, territory, and the District of Columbia archives organize, retain, collect, and manage state government social media records.

Toward a Common Understanding: Insights on Inter-agency State Electronic Records Transfer. DECEMBER 2019

This report documents the genesis, implementation, and findings of a research project targeted at shedding light on current practices associated with the inter-agency transfer of permanent electronic records in state governments.

Steps to Preserve State Electronic Records. 2019

Developed by CoSA and the National Association of Secretaries of State. This report was supported in part by a National Leadership Grant from the Institute of Museum and Library Services.

Top Tips for Government Agencies Working with Electronic Records. 2019

Developed by CoSA and the National Governors Association as part of a National Leadership Grant from the Institute of Museum and Library Services.

Managing Electronic Communications in Government. 2019

When is Information Misinformation? 2019

A collaborative publication of CoSA, the Chief Officers of State Library Agencies (COSLA), and the National Association of Secretaries of State (NASS), this handy infographic highlights common pitfalls for online users and offers five tips for avoiding them.

State Archiving in the Digital Era: A Playbook for the Preservation of Electronic Records. OCTOBER 2018

Developed by CoSA and the National Association of State Chief Information Officers, the playbook presents eleven “plays” to help state leaders think about the best ways to preserve archives in the digital era. This report was supported in part by a National Leadership Grant from the Institute of Museum and Library Services.

First Steps in Managing Records: A Transition Priority. 2018

Developed by CoSA and the National Governors Association, this document provides practical guidance for Executive Offices as they transition into and out of office. This brief was supported by a National Leadership Grant from the Institute of Museum and Library Services.

Social Media and Public Records: Developing a Social Media Use Policy. 2018

Developed by CoSA and the National Governors Association, this document provides archives and records management guidance on working with social media in government. This brief was supported by a National Leadership Grant from the Institute of Museum and Library Services.

A National Risk: The State of State Electronic Records Report. 2017

This report examines the impact of CoSA’s State Electronic Records Initiative (SERI) from its inception in 2011 while considering future directions for this program. This report was funded by a Laura Bush 21st Century Librarian Program grant from the Institute of Museum and Library Services.

These and other publications may be found at the CoSA website, <https://www.statearchivists.org/>



Council of State Archivists
PO Box 1583
Frankfort, KY 40602-1583

VOICE 502-229-8222
EMAIL info@statearchivists.org

Contents

Executive Summary	1
Introduction	2
Findings and Trends	3
Email Creators and IT in State and Territorial Governments	3
Archival Preservation and Practice	5
Obstacles to Email Preservation	7
Common Needs and Opportunities: The Path Forward	8
Email Policy Development and Adoption	8
Preservation Pathways for Email Records	8
Developing Discrete Projects to Increase Email Preservation Capacity	9
APPENDIX A	
PREPARE Needs Assessment Survey Questions	11
APPENDIX B	
PREPARE Advisory Group	19

Executive Summary

The Council of State Archivists (CoSA) is providing capacity-building services for email management and preservation to state and territorial archives through the University of Illinois' Email Archives: Building Capacity and Community program, funded by the [Andrew W. Mellon Foundation](#). In order to determine specific needs and interests, CoSA developed a needs assessment survey and distributed it to its member organizations in May 2021. There was a strong response from state and territorial archives—87.5% of the 56 states and territories completed the survey. These states shared their priorities and obstacles in email preservation including transfer, preservation, processing, access, institutional and state-wide policies, documentation, and resources. Responses to the survey and the insights they provide are discussed in aggregate in this report.

Email Preservation: Policy and Practice. Across the states, existing laws and regulations cover email as permanent public records requiring retention and preservation. Most states and territories have policies that cover both email and other electronic messaging as permanent records based on the content of the message rather than its format. This is necessary given the continual evolution of messaging technologies. However, despite more than 90% of responding states and territories reporting that existing laws, regulations,

and policies should result in the preservation of state government emails at the archives, permanent email records are inadequately preserved.

The PREPARE Needs Assessment survey demonstrates that email preservation at state and territorial archival agencies is still developing. Key figures include:

- 33% of respondents have not collected email records
- Fewer than 20% of responding state and territorial archives have been collecting email for longer than 10 years.
- Only 38% of respondents use their digital repository to preserve email records

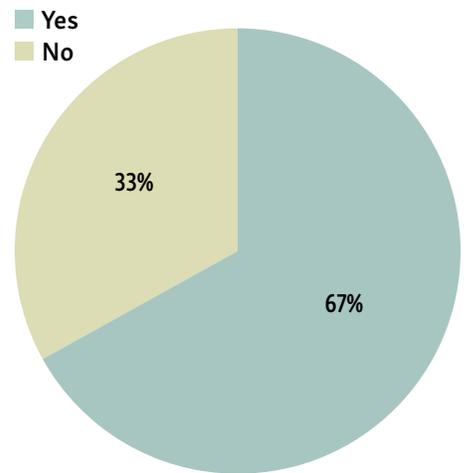
Challenges to State Government Email Preservation. Existing policies demonstrate that state and territorial governments are aware of the importance of email preservation. However, the lack of substantial collection and preservation of email records by state and territorial archives show that there are significant challenges implementing preservation practices.

Major impediments to preservation of email include:

- Huge scale of email records (66% of states listed this as a challenge and 19% identified it as the biggest obstacle they face)
- Insufficient statewide policy adoption (63% and 16%)
- A lack of access to technology (63% and 13%)

While states and territories find the scale of email records daunting now, the scale of email records is likely to continue trending upwards. A majority (82%) of responding archives expect to see their holdings of email grow

FIGURE 1. Does Your Archives Collect Email Records for Preservation?



significantly in the next 5 years. Yet 40% of responding archives report uncertainty about the specific amount of growth they will face.

State and territorial archives need to build additional capacity to successfully deal with the massive influx of email records and other electronic messaging records expected in the near future.

Closing the Gap: Opportunities to Increase Email Preservation.

Significant improvement in the preservation of email records could be made through collaborative guidance about the requirements of state and territorial email preservation practice. Legal requirements exist but are not widely implemented or understood. There is a strong opportunity for state and territorial executives to encourage adoption of policies that put agencies in alignment with legal mandates. Transfer of email records is a significant obstacle. Opportunities to improve content transfer include:

- Continued education about email records and the obligations of transferring them;
- Engagement and ongoing collaboration with state IT to support the transfer of data from content creator agencies to the state archives;
- Developing, publishing, and implementing state archive transfer expectations and workflows; and
- Integrating appropriate checks for data integrity and virus protection.

FROM THE SURVEY

5-year Email Growth Projections

- ▶ At minimum 1.4 million new email messages per archive
- ▶ Between 100GB–1TB of new email collections per archive
- ▶ Email records could easily account for more than 20% of state electronic records collections

Documentation is needed

- ▶ 58% of archives need requirements/specifications for email acquisition
- ▶ 71% of archives need file format specifications
- ▶ 90% of archives need metadata specifications
- ▶ 92% of archives need fixity specifications

Introduction

The Council of State Archivists (CoSA) is providing capacity-building services to state and territorial archives through the University of Illinois' Email Archives: Building Capacity and Community program, funded by the Andrew W. Mellon Foundation. CoSA PREPARE (Preparing Archives for Records in Email) seeks to increase capacity and capability for email preservation in state and territorial archives. In order to determine specific needs and interests, CoSA developed a needs assessment survey which was administered in May of 2021.

There was a strong response from state and territorial archives—roughly 88% of the 56 states and territories completed the survey. These states helped determine priorities and obstacles in email preservation including transfer, preservation, processing, access, institutional and state-wide policies, documentation, and resources. Responses to the survey and the insights they provide are discussed in aggregate in this report.

Furthermore, these results are contextualized by The State of State Archives biennial report. The report covers a broad range of data collected about the resources, authority, and condition of each state and territorial archives and records management programs with respect to electronic recordkeeping and digital preservation. Additionally, insights from the 2020 State Chief Information Officer (CIO) Survey by the National Association of State Chief Information Officers (NASCIO), a partner organization and close collaborator of CoSA, are included as well.

Findings and Trends

Findings from the PREPARE Needs Assessment Survey are organized into the following categories:

Email Creators and IT in State and Territorial Governments: While state and territorial archives operate under a legal mandate to manage permanent email records, these records require engagement from records creators and custodians, IT agencies, and other stakeholders as well. Some of the survey questions were posed to illuminate the practices of email management of non-archival stakeholders. For example, enterprise-wide information and data management policy (which is within the scope of authority of archives in some states, but not all) clarifies roles and expectations for email records management. This generally includes retention and preservation policies (e.g. records schedules). Although state and territorial archives have varying degrees of authority, influence, and control of scheduling processes, the preservation policy and practice of classification must be addressed.

Archival Preservation and Practice: This section details the management and preservation of email once it reaches the archives' authority, which is typically after retention timeframes have been met. Survey questions addressed both current and legacy practices, as well as aspirational goals. Specifically, many tasks related to the long-term preservation of email and other electronic messages will fall to the archives (and records management agencies). However, digital preservation requirements must necessarily be assigned to records custodians throughout the email record's lifecycle, much of which is outside the direct authority of the archives.

Obstacles to Email Preservation: Each state and territory has unique laws,

policies, IT governance, agency structures, archival authority, and dozens of other variables which contribute to the effective and efficient management and preservation of email. To cut through these variables, archivists were asked to identify those obstacles of greatest concern, as well as those of biggest impact.

EMAIL CREATORS AND IT IN STATE AND TERRITORIAL GOVERNMENTS

The PREPARE Project begins with some basic assumptions about the email preservation lifecycle:

- Email is created for business use in the course of public agency business
- It is housed in the custody of its creator(s) during its active life
- It is then stored in indefinite storage in its inactive phase while it awaits final disposition
- Whereupon it, having been classified appropriately by records schedules, will either be destroyed, or transferred to an appropriate repository for permanent preservation.

This reflects a simple model for the management of email that is similar

to the lifecycle of almost all electronic records. In this way, email is not unique in terms of its requirements.

The legal foundations necessary for the effective and efficient management of emails are sound; most states/territories have some degree of governance, either at the statutory or policy level, covering the retention and preservation of email. More than 90% of states reported that laws, regulations, and policies (including schedules) should result in emails being preserved at the archives. Additionally, close to 90% of states have records schedules that do cover or would cover email records.

However, there is some concern (27% of respondents) that records schedules are not up-to-date or sufficient enough to account for all nuances in dealing with email records; at the broadest level, the law is sufficient, but changes in practice and technology have created gaps (e.g. email being labeled as "routine correspondence" when it can be an important, permanent record). Despite retention schedules, statutes, laws and regulations that cover email records as part of the state or territorial archives purview, there is nevertheless a shortage of email records being transferred

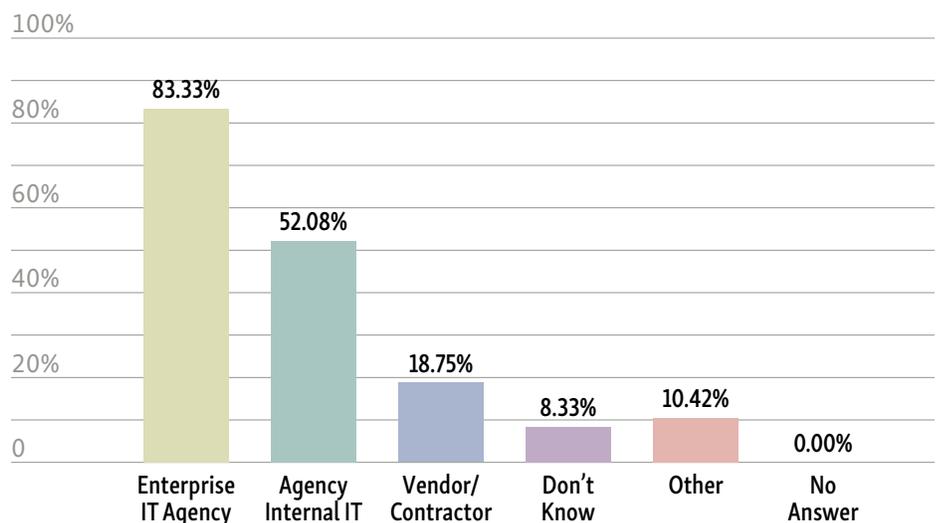


FIGURE 2. Who maintains and administers state email systems?

according to the existing and applicable records schedules or laws.

Other forms of electronic messaging do not have as robust policy coverage and will need additional attention in the future. Archivists in 37% responding states indicated that electronic messages are lacking applicable schedules or policies, or were unsure if existing policies covered these kinds of records. One gap that may exist is the specific classification of permanently-valuable information and the types of records wherein that information appears. While at its inception email may have been used primarily for inter-office communication, its uses now are more ubiquitous, and many kinds of essential evidence, including contextual information about government activity, is likely captured either within attachments or the emails themselves. A similar evolution of usage has occurred with other electronic messaging platforms, causing a gap in practice if not in policy.

However, despite good coverage in retention foundations, only 44% of reporting states have statewide policies that detail specific requirements for the preservation of email. CoSA notes that such policies are most likely to originate from the archives, records and information management (RIM), and information technology (IT). Classification and retention is only the beginning of the equation; email custodians still require training on the best methods to manage and preserve emails regardless of their classification.

 **ONLY 25% OF IT AGENCIES HAVE DISTRIBUTED STATE-WIDE EMAIL GUIDANCE.**

Despite email services being generally administered by either statewide IT agencies (in 83% of responding states) and/or agency IT (52%), only 25% of IT agencies have distributed state-wide email guidance. This is despite data from the most recent State CIO Survey, which registered that 68% of state IT agencies have distributed data governance policies. The lack of guidance regarding email records illustrates a divide between what IT might consider to be within its scope (e.g. technical

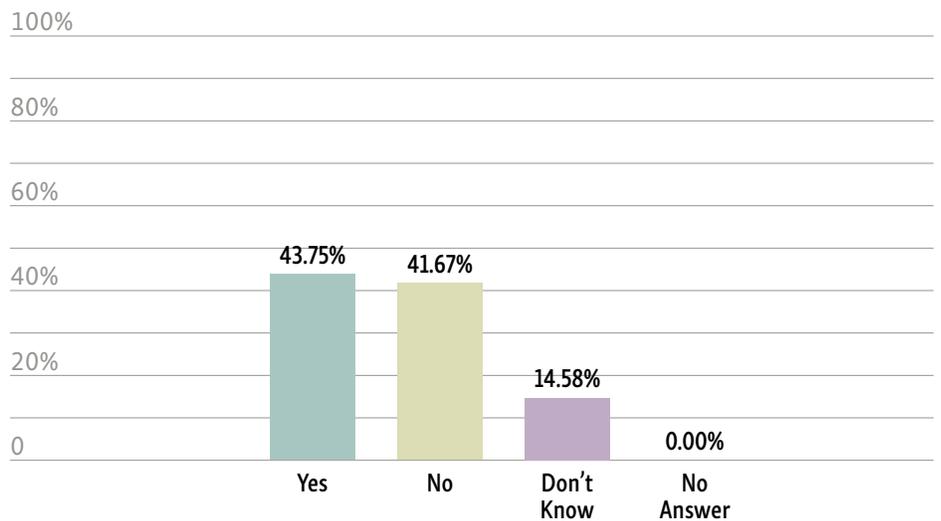


FIGURE 3. Do statewide policies exist that detail requirements for the management of email?

issues) and what it does not (e.g., retention issues). Without clear guidance and the assignment of responsibilities, the management of email records will lack consistency throughout the enterprise. Given the sheer volume of email expected to accumulate in the next half-decade, lack of rudimentary organization, metadata, and other information management basics is a growing concern.

Archives are trying to address this state-wide preservation policy gap independently; 21% of archives respondents have published their own email management requirements. In addition, a full third (33%) of state RIM programs have done the same. Much of the uncertainty around the establishment of state-wide

email requirements could be due to the unclear authority or the difficulty to enforce adoption of such policies. There is an opportunity to address this uncertainty by developing state-wide guidance that can be distributed by both the archives and state CIO offices. Of course, there are some states that have a more decentralized model for which this approach would not be possible.

The survey findings align with expectations and previously-gathered data about infrastructure. Only a few

respondents (4%) maintain email servers exclusively on state-controlled, on-premises servers; the vast majority of state-wide email systems are utilizing cloud-based email servers, or a combination of cloud and local servers. Despite the majority of state email largely stored on outside systems, it was surprising that only 19% of respondents have vendors or contractors involved in the management of state email. As more states are moving to consolidate their IT infrastructure (89% of respondents to the State CIO survey reported plans to expand “As-A-Service” models with 42% looking to downsize state owned and operated data centers), large cloud-based vendor platforms are very likely to take on a larger role in email delivery. Given the extensive usage of vendor-supplied technology, it is important to engage vendors as collaborators in producing both tools and outcomes that support long-term preservation of email data.

One encouraging set of responses concerns the usage of a Capstone approach to records management and preservation. About a quarter (25%) of respondents indicated that Capstone policies or approaches are formally present within their state or territory, even if they did not know the compliance or adoption of such policies. This could indicate, given that Capstone is ultimately a role-based approach to

appraisal and classification of email, that archives will be able to affect change on a broader level using role-based methods or approaches. Perhaps examining the practices used by these states to promulgate policies and training on Capstone (and other records management methodologies) could serve as models for others with regard to policy implementation.

Lastly, the needs survey examined the use of non-email electronic messages and the devices (and services) used in those contexts. More than 4 in 5 respondents (81%) indicated their state uses electronic chat protocols in some manner, while more than half (52%) reported that private messages or direct messages through social media platforms were common means of electronic communications. The challenges of electronic message preservation closely echo those of email, but are more complicated. Email, at least, has a fairly extensive body of standard protocols (like SMTP) and definitions which encourage interoperability. Few such standards exist for proprietary messaging services; archivists of the (near) future will need to tackle each messaging service individually, and given that the standard delivery model for electronic messages relies heavily

on software-as-a-service (SAAS), the records produced in this way are likely to be under vendor control where technology is hidden behind proprietary code.

All in all, the PREPARE responses present strong foundations that have yet to be capitalized upon, but also great opportunities for the improvement of email management and preservation



DESPITE RELATIVELY COMPREHENSIVE EMAIL POLICIES FOR RECORDS RETENTION, 33% OF STATE AND TERRITORIAL ARCHIVES HAVE NEVER COLLECTED EMAIL.

outside the archives. States are consolidating their infrastructure, and many have already taken steps to implement governance of email. Close collaboration between IT agencies looking to consolidate and archives looking to streamline the classification, transfer, and preservation of emails is highly suggested by the results of this survey.

ARCHIVAL PRESERVATION AND PRACTICE

Responses to the PREPARE Needs Assessment survey demonstrate that email preservation at state and territorial archival agencies is still in its infancy. Despite relatively comprehensive email

policies for records retention, 33% of state and territorial archives have never collected email. Of the 67% of state and territorial archives who responded that they have and are collecting emails, only 19% have been doing it for more than a decade. Thus, close to 81% of archives have been collecting email for less than 10 years, despite it being used in state

government for decades. While many state archives have been involved in the management of email in some shape or form, until recently it has not risen to the level of a formal policy or of transfer and deposit in an archival repository.

Selection of records through classification in schedules, or the application of those taxonomies does not seem to be the problem, given that more than 90% of respondents indicated that email should be preserved given the current state of laws, regulations, schedules, and other parts of public policy. There is, or should be, considerable overlap between archival practice and the management of records throughout the entire lifecycle. For decades, many emails have been undergoing “preservation in place” (being stored in email servers or electronic records management systems until appropriate action can be taken), and are at risk of being lost. Attempting long-term preservation in systems not designed for such tasks is risky at best, and unsustainable in the long-term.

The primary difficulty for the archives is simply getting email records transferred to their custody. Transfer is a key focus for the next five years: 29% of respondents identify it as a priority. Lessons from the MoVE-IT project have indicated that transfer of permanent electronic records is already an area of great concern, and PREPARE confirms that email is equally susceptible to content transfer obstacles. The lack of policy and expertise across the enterprise in this area may be making the execution of transfer difficult; indeed, while many states see at least some

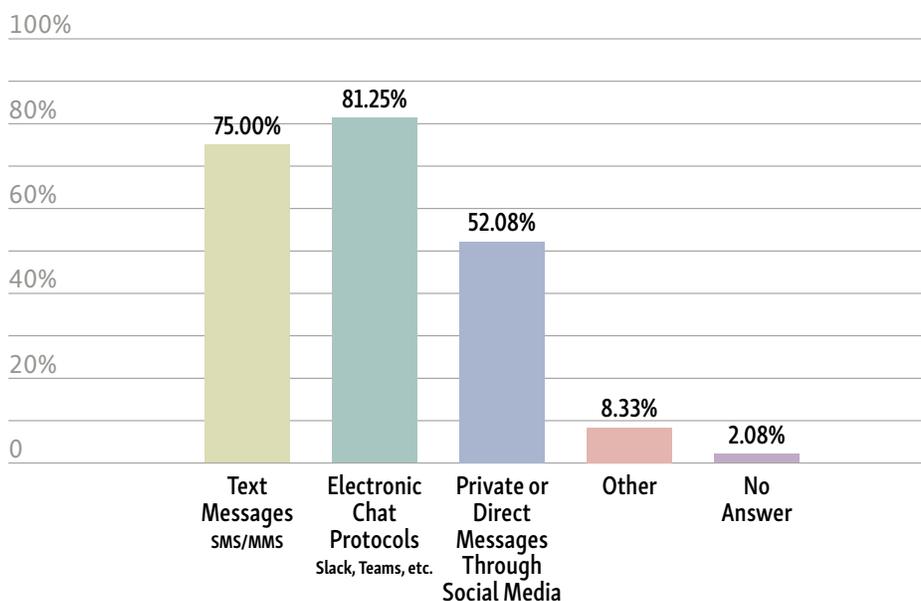


FIGURE 4. What kinds of non-email electronic messages are currently being used across government for official business?

email transfers conducted by IT agencies (63%), archives in many states see regular transfers by RIM staff, vendors, legal departments, and others. There are few standardized approaches to transferring records and archives are currently using a “get them any way you can” model. And while an IT agency might be the most appropriate agency to conduct the data transfer given their likely administration of email servers or relationship with vendors who do so, classification and selection of records is a task likely better suited to subject-matter experts. While individual states may have defined transfer pathways, there is little consistency surrounding email transfer to archives across the United States. Definition of requirements, tasks, and roles as part of these pathways can assist archives in the assignment of responsibilities throughout the lifecycle.

While merely getting email records to the archives is a major obstacle, it is magnified by the expected scale and growth of email records. A majority (51%) of archives expect to see their holdings of email grow by at least 100GB in the next 5 years, and many predict even more growth. At an average size of roughly 75kb per email, a 100GB growth equates to nearly 1.4 million new messages). Other archives (31%) expect their email collections to grow even more, by at least 1TB or more. There is great uncertainty in these projections, as 40% are unsure of how much their holdings could grow. This is understandable, especially considering that, at the end of FY 2020, the median volume of electronic records in state and territorial archival holdings was around 5TB; this would mean that email could reasonably be expected to account for up to 20% or more of state electronic records collections in the next half-decade, and archives are not adequately equipped to deal with this massive expected influx of electronic records.

It is possible that lack of guidance from the archives may be constraining state government. Nearly 3 out of 5 archives (58%) have no requirements or specifications for the acquisition of email; that is, records creators and IT

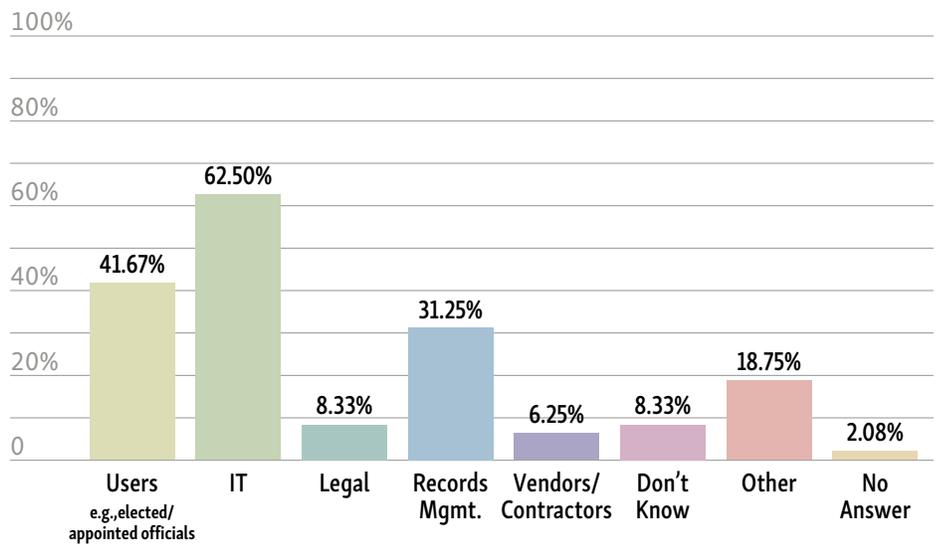


FIGURE 5. Who typically performs (or would perform) a transfer of email records to the archives?

agencies need do nothing more than transfer emails “as-is” to the archives for permanent preservation. While some archives have implemented file format specifications (29%), metadata specifications (10%), and technical specifications (8% for fixity and 12% for other), it’s unlikely that specifications are impeding transfers of email. More likely, even where specifications exist (either for records or for process), there is substantial difficulty across the enterprise to ensure policy adoption (63% reported this being an obstacle). It is possible that lack of formal guidance presents records custodians with little idea of how to approach a management problem like email.

Even when emails do make their way to the archives, the available technological solutions may not be adequate to handle the challenges of processing, preserving, and presenting these

records. Only 49% of state/territorial archives have an OAIS-compliant digital preservation repository (though many use electronic records management systems (29%), external hard drives (45%), shared network storage (55%) and other methodologies for hosting data for permanent preservation).¹ Yet, despite nearly half of respondents having a repository for the purpose of preserving electronic records, only 38% of respondents use their digital repository to preserve email records. This could indicate some incompatibilities between commonly-used OAIS repositories and email preservation requirements, or could be a matter of scale (volume of email in digital repositories could cause the size of collections and hosting costs to balloon).

More concerning is the lack of reach of email processing tools created by and for the archival community. There has

5-year Email Growth Projections

- ▶ At minimum 1.4 million new email messages per archive
- ▶ Between 100GB–1TB of new email collections per archive
- ▶ Email records could easily account for more than 20% of state electronic records collections

Documentation is needed

- ▶ 58% of archives need requirements/specifications for email acquisition
- ▶ 71% of archives need file format specifications
- ▶ 90% of archives need metadata specifications
- ▶ 92% of archives need fixity specifications

FROM THE SURVEY

¹ CoSA. [The State of State Records](#). 2020.

been a lack of penetration of tools like DArMail,¹ TOMES,² RATOM,³ ePADD,⁴ each of which is used in less than 4% of responding organizations. State and territorial archives sometimes lack adequate IT support while facing IT security concerns that make open-source software difficult to get through security reviews. Such tools are desperately needed by state and territorial archives, as archives are struggling to provide common digital preservation tasks such as surveying material for exempt and otherwise protected information (conducted by archives staff



ALTHOUGH THE LAW SUPPORTS ARCHIVAL PRIORITIES, IN PRACTICE, EMAIL RECORDS AREN'T MAKING IT TO THE ARCHIVES.

at 60% of respondents), file format requirements (29%), preserving attachments (23% not preserved) and other email-adjacent records (e.g. links to documents, attachments, etc.), and specifying metadata (10%), fixity (8%). Lack of processing tools, or expertise with them (42% responded that archival staff require more training), is contributing to backlogs.

Archives that are successful in getting permanent email records transferred to their custody still face challenges and obstacles to managing those permanent records. Archives generally do not have experience or infrastructure with specialized software to process email, most often due to the investment required to do so. Furthermore, archives aren't using their digital repositories to preserve their email collections as much as they could. As with the data collected about enterprise policy and process, these data points present opportunities for state and territorial archives to increase their capacity.

OBSTACLES TO EMAIL PRESERVATION

Finally, states and territories were asked to discuss their biggest obstacles to efficient and effective preservation of email across the enterprise. The challenges of greatest concern include:

- **Technology:** State and territorial archives lack the technology they need to manage email preservation. This includes digital repositories that can support email preservation requirements as well as software to process, analyze, and describe permanent email records collections.
- **Standardized Models:** While states and territories have statutes and laws that identify the content of email as permanent records, statewide policies aren't universally adopted by records-creating agencies. Although the law supports archival priorities, in practice, email records aren't making it to the archives as much as would be expected. Developing standardized models and templates for preservation activities including policies, expectations, roles, and responsibilities is a key first step.
- **Scale:** The overwhelming volume of email is a scale issue that archives are struggling to address. Email sets require significant investments in order to properly classify, select, and describe large quantities of records. Unlike some records originating from other programs, accumulation of email is continuous and expanding.
- **Policy Documents:** State and territorial archives lack written policy for managing key digital lifecycle phases.

Archives indicate a significant interest in developing strong email management policies as well as managing the preservation activities required to ensure the permanence of valuable email records.

- **Training/Experience:** Archives' staff do not feel adequately trained to manage email preservation—and by extension to manage digital preservation. 42% believe the archives need additional training to preserve email records, while 50% report they believe records creators need additional training. Email and electronic messaging technology is continually changing, so training and education programs must also be continuous to keep up with the technological landscape of state government.

- **Content transfer:** The five-year goals for progress in email preservation show that there is significant work to be accomplished regarding transfer of state email records to permanent repositories.

Note: Description of and access to email records were not identified as major priorities in the five-year goals for state and territorial archives. This does not indicate lack of interest but rather a more pressing need to address issues earlier in the life cycle. Since records rarely make it to the archives in a state conducive to access, providing access is currently a distant ideal rather than an achievable goal. Should progress be made in transfer, acquisition, processing, and preservation, the areas of description and access would be the next categories of focus for state and territorial archives.

¹ <https://siarchives.si.edu/what-we-do/digital-curation/email-preservation-darcmal>

² <https://github.com/StateArchivesOfNorthCarolina/ratom-server>

³ <https://ratom.web.unc.edu/tools-and-code/>

⁴ <https://library.stanford.edu/projects/epadd>

Common Needs and Opportunities: The Path Forward

EMAIL POLICY DEVELOPMENT AND ADOPTION

As discussed, 90% of states and territories have legal mandates regarding permanent email records. The adoption and implementation of policies and procedures to actualize such mandates, however, has not yet been realized. Archivists have identified policy adoption as a major priority in the next 5 years. Successfully implementing policies based on industry standards and best practices will be essential to realizing state-wide email preservation throughout the electronic records lifecycle.

Email management and preservation policy can vary greatly from state to state. Each state is unique in their situation with regard to both archival influence and authority. However, email lifecycle management is very likely to involve multiple agency actors regardless of the position of the archives. All state archives should be, if they are not already, experts in digital preservation, and policy development in this area must begin where the records lifecycle begins, at the point of creation. One best-practices and standards-oriented approach to enterprise email management could be broadly applicable to every jurisdiction. This model would focus on individual tasks and services to be performed to ensure emails are properly preserved at every point of their lifecycle, and that whoever needs to access them (the creator, IT, legal, HR, the public) has the proper, legal channels to do so.

Since the archives has little influence over the day-to-day workflows of government employees, policy adoption will need to come in collaboration with records-creating agencies. Archives may have more influence over information

systems, but collaboration with state IT agencies and staff will ensure that preservation requirements are embedded into procurement processes and technology contracts. CoSA must continue to work with institutional partners like the National Association of State Chief Information Officers to foster collaborations and enhance relationship-building in individual states. A collaboration among records creators, IT, and archivists must be present in each state in order to advance digital preservation policy for email and other electronic messages.

PRESERVATION PATHWAYS FOR EMAIL RECORDS

Email in state government is plagued by two major issues: 1) despite legal mandates, email is not regularly transferred to the archives, and 2) when it is transferred, it is usually unprocessed,¹ placing a huge burden on the archives' ability to provide access to such records.



ONE POSSIBLE PATH FORWARD WOULD INVOLVE COSA WORKING WITH STATE ARCHIVES TO DEVELOP STANDARDIZED "PRESERVATION PATHWAYS" FOR EMAIL.

Email housed in business systems unequipped for digital preservation is a major risk to state government. Additionally, email requires processing (especially surveying for protected information such as PII or PHI), which is especially burdensome given the resource-intensive nature of such work. The typical electronic record lifecycle should result in records being destroyed or transferred when they reach final disposition, but for email the process is breaking down, likely due to one of three factors:

- Records creators are unable to identify permanent records or are unable/unwilling² to transfer them;
- Records creators can identify and wish to transfer records, but there is no mechanism to do so (or the mechanism is cumbersome or resource-intensive); or
- Records are transferred to the archives (or are ready for transfer), but the archives are unable to accept/process/preserve/provide for the records appropriately.

The obstacles can also be present in any or all of these areas, whether that obstacle is a lack of policy, a lack of adoption, a lack of compliance once adopted, a lack of adequate training, or lack of staff time devoted to processing. While there are many different models of state government, it's likely that there are few truly different email delivery models. Understanding these models can allow PREPARE to dissect, analyze, and work with CoSA's community to develop best practices tailored to each of these models. Comprehensive

understanding of email solutions architecture will enable the archival community to best collaborate with its major stakeholders with a basis of common understanding.

One possible path forward would involve CoSA working with state archives to develop standardized "preservation pathways" for email. Email requires considerable lifecycle management and specialized tools (like those created as part of the RATOM, TOMES, and ePADD projects) to manage it, in part because of the scale and method

¹ Meaning "requiring arrangement, description, and/or survey for protected information"

² Unwillingness includes the inability to do so for budget/resource reasons

of organization. By defining several standard email delivery, classification, and transfer models, states could customize these basic frameworks to match their individual circumstances. There is an opportunity to work with IT departments to prime agencies, IT, and archives for new tool adoption that will build bridges between records creators and preservation agencies. Advocacy for records management to IT departments is a message that bears repeating.

Preservation lifecycle elements for email and electronic messaging records must include:

- **Classification and Taxonomy:** Simple yet thorough classification and metadata standards to serve as the data framework for email lifecycle management. Roles and responsibilities related to classification of email will likely involve everyday users and thus should be achievable with minimal training.

- **Archival Appraisal and Processing:** Policy (Capstone) and/or technological (artificial intelligence) approaches to identification and selection of records. Processing to identify Personal Identifiable Information (PII), to restrict or limit usage as legally mandated, and to deselect emails that are not permanent records.

- **Transfer:** Physical and intellectual exchange of record custody, preferably into digital preservation repositories.

- **Storage and Preservation:** Identification of both short-term and long-term preservation requirements for electronic records and repositories.

- **Search and Retrieval:** Discovery of and access to email records for concerns both internal and external to state government.

- **Migration:** Email will need to be moved into new formats as well as new storage media to ensure preservation and access for records and repositories of permanent value.

These roles and responsibilities can be performed by any combination of state government agencies in ways that support the capabilities and needs of the state or territory.

DEVELOPING DISCRETE PROJECTS TO INCREASE EMAIL PRESERVATION CAPACITY

A path forward requires strategically discrete, actionable steps. Increasing state and territorial archival capacity to preserve email and electronic messages is a requirement for the successful implementation of an institution's commitment to long-term email records preservation. This commitment must extend beyond the archives and integrate fully into the enterprise technology environment of the entire state or territory, supported by the laws and policies that govern public records and their administration.

CoSA recommends that state and territorial archives attack specific email preservation problems in systematic ways, thereby developing solutions that can be built upon to create a comprehensive email preservation solution. Building elements of a full email preservation lifecycle solution piece by piece to demonstrate success and garner support will help archives establish their long-term preservation goals across the entire lifecycle.

When creating new preservation projects, seeking additional funding through the state budget, or writing grants to support new preservation activities, CoSA recommends that archives focus strategically on capacity-building projects. Clear and precise standards and procedures specifying selection of records for preservation, file format preferences, transfer practices, metadata schemas, fixity, or other technical specifications communicate clearly what resources are needed and that email preservation is a shared responsibility. Some examples of discrete documentation projects that support institutional preservation include:

- Formalize state or territorial archives transfer policy and practices in a comprehensive set of procedures. If the state uses a Capstone approach, or if records creators have specific responsibilities under the legal mandates of their

state or territory, states should consider creating a manual that describes these requirements both concisely for easy adoption, and comprehensively for a full analysis. Distribute and discuss with as many email-creating agencies and individuals as possible.

- Developing easy-to-consume records management materials such as pre-recorded webinars or short videos that answer specific questions for email creators.

- Collaborate with the state's Chief Information Officer (CIO) and/or senior level state IT staff as potential partners to develop and disseminate guidelines on applying email records management requirements to content creators.

- Through the PREPARE Project, CoSA can foster the development of policy by assembling best practices and techniques for tasks and requirements throughout the lifecycle of email.

- Formalize the archives' preservation lifecycle policies. If there is no clear documentation regarding transfer of electronic records, develop clear transfer procedures:¹

- Guidance like the MoVE-IT report can provide ideas about how to make content transfers successful in the state or territory.

- Use CoSA Format Guidelines to inform requirements about the preferred formats for email preservation. Integrating selection, arrangement, and description standards into lifecycle management guidelines can alleviate some of the processing burden currently falling nearly exclusively on the archives.

- Work with the state CIO's office, state agency IT staff, and state legal counsel to determine what support they can offer for identification and selection of emails for transfer, and for automating said transfer of electronic records to the archives. In states and territories where there is an enterprise email solution, the state CIO's office can help bridge the gap between the laws, statutes, regulations, and policies that support the preservation of email records and the lack of content getting to archives. Archives

¹ Consider using the tools created as part of the CoSA MoVE-IT Project to facilitate the development of transfer workflows and procedures: <https://www.statearchivists.org/programs/state-electronic-records-initiative/move-it-modeling-viable-electronic-information-transfers/>

that have clear requests with specific steps around individual issues are more likely to get attention and support from state CIO's offices. Informed staff are more likely to be successful in any collaboration with organizational stakeholders.

CoSA is dedicated to supporting the work of state and territorial archives in all endeavors, especially digital preservation. CoSA's commitment to research,

advocacy, and community position it to support state and territorial archives in their work to increase email transfer and preservation capacity. The PREPARE Needs Assessment Survey reaffirms CoSA's focus on creating and sharing research-driven best practices, developing templates and guidelines to assist archives, developing policies and procedures based on industry standards,

supporting communities of practice, and furthering collaborative partnerships.

Information on the work CoSA is doing to support increased capacity and capability in state and territorial archives for email and electronic messaging, as well as all electronic state records, can be found at <https://www.statearchivists.org/programs/state-electronic-records-initiative/>.

PREPARE Needs Assessment Survey Questions

The PREPARE Needs Assessment Survey was developed over the course of several months by the PREPARE staff, with direct consultation from an advisory group made up of individuals representing state, federal, and university archives, the IT sector, as well as not-for-profit organizations. A full list of the PREPARE advisory group can be found in Appendix B.

Initially, the Project Manager drafted a list of questions aiming to address broad issues (and in some cases, specific details) pertaining to the entire lifecycle of email creation, use, and preservation. This set of roughly 50 questions was then reviewed by the project steering committee and ultimately SERI leadership before being shared with the PREPARE Advisory Group for consultation and review. At the suggestion of the advisory group, other kinds of electronic messages, such as instant messaging, group chat services, and others were considered, but ultimately were removed from the survey to increase clarity in the data, except in select instances.

Once the survey questions were finalized, it was shared directly with State and Territorial Archivists, who were given 3 weeks (later extended to 5 weeks) in order to complete and return the survey to CoSA. All in all, 49 responses were received.

SOME DEFINITIONS

Email: Refers to electronic text messages and attachments as defined in IETF RFC 5322.

Electronic Messages: Non-email messages such as: text messages (MMS or SMS), direct messages through social media platforms, instant message programs, chat, and so on.

State: Use of “state” in this survey includes the entirety of state or territorial government.

Archives: The legally-designated state or territorial archival agency, regardless of where it is located within the frame of government.

SURVEY

The CoSA PREPARE Project seeks to ascertain the degree of organizational context, electronic recordkeeping maturity, and preparedness related to the description, transfer, preservation, and access to email and other electronic messages present in states, territories, and their archival institutions across the U.S. The goal of the PREPARE project is to establish a community of practice around the successful and efficient management and preservation of email. This survey will allow CoSA to determine how to best use our resources and what outputs will be most beneficial to the state and territorial archival community.

Please fill out the questions below to the best of your knowledge; it is possible that some of this information will not be available to the archives or its staff. It is valuable to us to know what information is unavailable. Please report that, too.

The survey should take roughly an hour to fill out and may require some additional research. Please note that you are able to save your progress and return to the survey, so it does not need to be completed in one sitting.

Thank you in advance for taking the time to fill out this survey for the Council of State Archivists! For more information on this project, please visit the project website: <https://www.statearchivists.org/programs/state-electronic-records-initiative/cosa-prepare-preparing-archives-records-email/>

1 Please state the full name of the institution you represent:

2 What jurisdiction does this institution represent? (Select one option)

- | | | | |
|--|-------------------------------------|--|---|
| <input type="radio"/> Alabama | <input type="radio"/> Illinois | <input type="radio"/> Nevada | <input type="radio"/> South Carolina |
| <input type="radio"/> Alaska | <input type="radio"/> Indiana | <input type="radio"/> New Hampshire | <input type="radio"/> South Dakota |
| <input type="radio"/> American Samoa | <input type="radio"/> Iowa | <input type="radio"/> New Jersey | <input type="radio"/> Tennessee |
| <input type="radio"/> Arizona | <input type="radio"/> Kansas | <input type="radio"/> New Mexico | <input type="radio"/> Texas |
| <input type="radio"/> Arkansas | <input type="radio"/> Kentucky | <input type="radio"/> New York | <input type="radio"/> Utah |
| <input type="radio"/> California | <input type="radio"/> Louisiana | <input type="radio"/> North Carolina | <input type="radio"/> U.S. Virgin Islands |
| <input type="radio"/> Colorado | <input type="radio"/> Maine | <input type="radio"/> North Dakota | <input type="radio"/> Vermont |
| <input type="radio"/> Connecticut | <input type="radio"/> Maryland | <input type="radio"/> Northern Mariana Islands | <input type="radio"/> Virginia |
| <input type="radio"/> Delaware | <input type="radio"/> Massachusetts | <input type="radio"/> Ohio | <input type="radio"/> Washington |
| <input type="radio"/> District of Columbia | <input type="radio"/> Michigan | <input type="radio"/> Oklahoma | <input type="radio"/> West Virginia |
| <input type="radio"/> Florida | <input type="radio"/> Minnesota | <input type="radio"/> Oregon | <input type="radio"/> Wisconsin |
| <input type="radio"/> Georgia | <input type="radio"/> Mississippi | <input type="radio"/> Pennsylvania | <input type="radio"/> Wyoming |
| <input type="radio"/> Guam | <input type="radio"/> Missouri | <input type="radio"/> Puerto Rico | |
| <input type="radio"/> Hawaii | <input type="radio"/> Montana | <input type="radio"/> Rhode Island | |
| <input type="radio"/> Idaho | <input type="radio"/> Nebraska | | |

3 In your state, is the retention of email covered by existing records retention schedules? (Select one option)

- Yes
- No
- I don't know

4 Are electronic messages (texts, instant messages, online chat, and other non-email messages) covered by existing records retention schedules? (Select one option)

- Yes
- No
- I don't know

5 Do statewide policies exist that detail requirements for the management of email? (include automatic-deletion policies) (Select one option)

- Yes
- No
- I don't know

NOTE: Answer the below question only if answer to **Question 5** is Yes.

6 If there are policies, from where do those policies originate? (Check all that apply)

- Archives
- Records Management
- Human Resources
- IT
- Other (Please specify): _____

NOTE : Answer the below question only if answer to **Question 5** is Yes

7 If possible, describe the nature of these policies:

8 What kinds of non-email electronic messages are currently being used across government for official business? (Check all that apply)

- Text messages (SMS/MMS)
- Electronic chat protocols (Slack, Teams, etc.)
- Private or Direct Messages through social media
- Other (Please specify): _____

9 In the state/territory, is there an agency responsible for managing email and other electronic communications services? (Select one option)

- Yes
- No (agencies manage their own systems)
- I don't know

10 Who maintains and administers state email systems (Check all that apply)?

- Enterprise IT Agency
- Agency Internal IT
- Vendor / Contractor
- I don't know
- Other (Please specify): _____

11 Has any agency (including the state/territorial archives agency) in your state adopted a version of NARA's Capstone approach for preserving email records? (Under Capstone, records in the email accounts of designated senior officials are deemed permanent, and all other employees' program-related email records are preserved for seven years under a General Records Schedule.) (Select one option)

- Yes
- No
- I don't know

NOTE: Answer the below question only if answer to **Question 11** is Yes.

12 If yes (Capstone), how many agencies (by percentage) have adopted the policy? (Select one option)

- 1-25%
- 26-50%
- 51-75%
- 76-100%
- Not sure how many

13 Are the tools used in the creation and management of email hosted in on-premises state servers or cloud-hosted servers?

- On-Premises
- Cloud
- Both
- I don't know

14 Do agencies in your state/territory have "bring your own device policies"? What estimated percentage of agencies have such policies? (Select one option)

- 0%
- 1 to 25%
- 25 to 50%
- 50 to 75%
- 75 to 100%
- Yes, but don't know percentage
- Enterprise-wide policy
- I don't know

15 Does the state/territory require that electronic messaging for official business be performed through state/territorial messaging applications only? (Select one option)

- Yes
- No
- I don't know

NOTE: Answer the below question only if answer to Question 15 is Yes.

16 What is the estimated compliance level with such policies? (Select one option)

- 0%
- 1 to 25%
- 25 to 50%
- 50 to 75%
- 75 to 100%
- I don't know

17 Does the archives currently (or was it previously) collect, receive, or accession email? (Select one option)

- Yes
- No
- I don't know

NOTE: Answer the below question only if answer to Q#17 is Yes.

18 How long has the archives been collecting email? (Select one option)

- n/a
- 1-5 years
- 5-10 years
- 10-20 years
- 20+ years

NOTE: Answer the below question only if answer to Question 17 is Yes.

19 What is the total volume of electronic messages your institution has collected (by number of electronic messages)? (Estimates are acceptable) (Select one option)

- 1M
- 1M-5M
- 6M-10M
- 10M-25M
- 25M+
- I don't know

NOTE: Answer the below question only if answer to Question 17 is Yes.

20 What is the total volume of electronic messages your institution has collected (by size)? (Estimates are acceptable) (Select one option)

- <1GB
- 1GB-10GB
- 10GB-100GB
- 100GB-1TB
- 1TB-10TB
- >10TB
- I don't know

21 How much do you expect the total size of your collected electronic message holdings to increase in the next 5 years? (Select one option)

- 1GB-10GB
- 10GB-100GB
- 100GB-1TB
- 1TB-10TB
- >10TB
- I don't know

22 Whose email are or would be collected under current laws, regulations, schedules, or other policies? (Please check all that apply)

- Elected Officials
- Executive Branch Agencies
- Judicial Branch
- Legislative Branch
- Non-Public Creators (Organizations and Individuals)
- General Public

NOTE: Answer the below question only if answer to **Question 17** is Yes.

23 Does the archives acquire emails and other electronic messages directly from servers, or from clients (e.g. employees mailboxes)? (Select one option)

- Servers (IT / Enterprise Platform)
- Clients (Individual Users)
- Both
- I don't know
- Other (Please specify): _____

24 Who typically performs (or would perform) a transfer of email records to the archives?

- Users (e.g. elected/appointed officials, program staff)
- IT
- Legal
- Records Management
- Vendors/Contractors
- I don't know
- Other (Please specify): _____

25 Does the archives acquire/collect individual messages or entire mailboxes? (Select one option)

- Messages
- Mailboxes
- Both

26 Does the archives have requirements or specifications for the acceptance or acquisition of email or other electronic message records? (Check all that apply)

- No specifications
- File format / characterization specifications
- Metadata specifications
- Fixity specifications
- Other (Please specify): _____

27 Does the archives accept or require the submission of reformatted electronic messages (printed, converted to open formats, etc.)? (Select one option)

- Accepts native originals
- Accepts non-archives reformats
- Reformats/Normalizes after accession
- I don't know

28 What kinds of electronic messages are currently being collected, acquired, and preserved by the archives? (Check all that apply)

- Text messages (SMS/MMS)
- Electronic chat protocols (Slack, Teams, etc.)
- Private or Direct Messages through social media
- Other (Please specify): _____

29 What file formats does the archives acquire/accept with regard to email and other electronic messaging services?

- | | |
|---|---------------------------------------|
| <input type="radio"/> .pdf | <input type="radio"/> .cca (obsolete) |
| <input type="radio"/> .eml | <input type="radio"/> SMS |
| <input type="radio"/> .msg | <input type="radio"/> NSF |
| <input type="radio"/> .pst | <input type="radio"/> IMF |
| <input type="radio"/> .txt | <input type="radio"/> XML |
| <input type="radio"/> .mbox family | <input type="radio"/> No limitations |
| <input type="radio"/> Other (Please specify): _____ | |

30 What metadata does the archives regularly capture for email messages (Check all that apply)

- | | |
|---|---|
| <input type="radio"/> From | <input type="radio"/> Dates |
| <input type="radio"/> Sender | <input type="radio"/> Content/Structure |
| <input type="radio"/> Reply-To | <input type="radio"/> Comments |
| <input type="radio"/> To | <input type="radio"/> Keywords |
| <input type="radio"/> CC | <input type="radio"/> Flagged |
| <input type="radio"/> BCC | <input type="radio"/> Category |
| <input type="radio"/> Subject | <input type="radio"/> Sensitivity |
| <input type="radio"/> Other (Please specify): _____ | |

31 How are attachments preserved, if they are preserved? (Select one option)

- In mailbox
- In message
- Separately
- Separate but contextually linked
- Reformatted/Normalized
- Not preserved / Email not currently preserved
- Other (Please specify): _____

32 How are other email-related items preserved (including calendars, contacts, tasks, notes, journals, etc.)? (Select one option)

- With mailbox
- Separately
- Separate but contextually linked
- Reformatted/Normalized
- Not preserved / Email not currently preserved
- Other (Please specify): _____

33 Do you have a method to uniquely identify senders and recipients? (email address, digital signature, etc.) (Select one option)

- Yes
- No
- I don't know

NOTE: Answer the below question only if answer to **Question 33** is Yes.

34 Please describe the method for identifying agents (senders/recipients):

35 What software is used to process and/or preserve email records? (Check all that apply)

- Digital Repository
- Database
- e-Discovery
- BitCurator
- DArcMail
- Other (Please specify): _____
- TOMES
- RATOM
- ePADD
- Emailchemy

36 Who (if any) is responsible for the review of retrieved messages for potentially exempt or non-public material?

- Employees/End-Users
- Records Management
- IT
- Legal
- Administrative staff
- Archives
- Other (Please specify): _____

37 How does the archives provide search and retrieval of email records?

- Finding Aid/Index
- Full-text search
- Sender/Recipient search
- Metadata search
- Browse
- Search not available at this time
- Other (Please specify): _____

38 How does the archives provide search and retrieval of email records? (Check all that apply)

- Finding Aid/Index
- Full-text search
- Sender/Recipient search
- Metadata search
- Browse
- No access currently available
- Other (Please specify): _____

39 If there are limited or no email or other electronic message policies, what is the greatest impediment to adoption of policies and procedures? (Select one option)

- Administrative Priorities
- Difficulty of adoption
- Execution (limited staff)
- Execution (limited budget)
- Other (Please specify): _____

40 Who is responsible for search and retrieval of electronic messages (including FOIA and other public records requests) during the active part of their lifecycle?

- Employees/End-Users
- Records Management
- IT
- Legal
- Administrative staff
- Archives
- I don't know
- Other (Please specify): _____

41 What obstacles exist to the efficient and effective preservation of electronic messages in your state/territory? (Check all that apply)

- Archives lacks necessary technology
- Agencies lack necessary technology
- Archives lacks necessary training
- Agencies lack necessary training
- Costs too high
- Challenges with statewide policy adoption
- Archives lacks necessary authority
- Record schedules not up-to-date / sufficient
- Privacy Concerns
- Legal Issues
- Scale/Quantity of Email
- Incompatibility of Policies/Procedures
- Lack of Agency Cooperation
- Lack of IT Cooperation
- Other (Please specify): _____

42 Which of these would you say is the biggest obstacle? (Select one option)

- Archives lacks necessary technology
- Agencies lack necessary technology
- Archives lacks necessary training
- Agencies lack necessary training
- Costs too high
- Challenges with statewide policy adoption
- Archives lacks necessary authority
- Record schedules not up-to-date/sufficient
- Privacy Concerns
- Legal Issues
- Scale/Quantity of Email
- Incompatibility of Policies/Procedures
- Lack of Agency Cooperation
- Lack of IT Cooperation
- Other (Please specify): _____

43 In terms of organizational maturity, where would the archives like to make the most progress in managing electronic messages (in the next 5 years): (Select one option)

- Email Management Policy
- Transfer
- Appraisal/Processing
- Preservation
- Description
- Access
- Other (Please specify): _____

44 How can the Council of State Archivists best assist your state's management and preservation of email? (Select one option)

- Development of Policy/Procedure
- Highlighting/Sharing Best Practices
- Testing and Reporting on Tools
- Developing Collaborative Partnerships
- Building Communities of Practice
- Other (Please specify): _____

45 In what ways can the Council of State Archivists assist in development of policy or procedure related to the preservation and management of email and other electronic messaging records?

As part of the PREPARE Project, CoSA is seeking email data sets to be used for testing suites of tools and procedures. CoSA is looking for any collection of email records, regardless of size or characteristics, to use for testing purposes. While we prefer raw, unprocessed sets of email (accounts and messages), any will do.

The work of the PREPARE project will involve using email management and processing tools on these data sets, including those that search for protected information such as PII and PHI, in tandem with expert archivists from its membership. CoSA will work with you to ensure that any concerns as well as all applicable state/territorial laws, regulations, policies, and requirements are respected should you choose to share records for this project.

We hope that “lending” email in this way can provide a jump-start to email processing and management at your institution, and we would love to work with archives who even to this point have done little to no work with email, as long as email records are available.

46 Does your state/territory have sets of email it would be willing to share with the CoSA PREPARE Project? (Select one option)

- Yes
- Maybe
- No

NOTE: Answer the below question only if answer to **Question 46** is Yes OR Maybe

47 Can you please describe the nature of these sets of email?

NOTE: Answer the below question only if answer to **Question 46** is Yes OR Maybe.

48 Can you provide the name, title, and contact information of a staff member whom we could contact about potential use of these records?

PREPARE Advisory Group

The work of PREPARE is improved and guided by the thoughtful participation of the PREPARE advisory group. Our sincere thanks to:

Jason R. Baron, University of Maryland
Roger Christman, Library of Virginia
Stephanie Clark, Oregon State Archives
Bryan Collars, South Carolina State Archives
Glynn Edwards, Stanford University
Deborah Gayle, National Archives and Records Administration
Joanne Kaczmarek, University of Illinois
Julio Lopez, Gates Archive
Kathleen O'Neill, Library of Congress
Elizabeth Perkes, Utah State Archives
Jeanette Plante, U.S. Department of Justice
Camille Tyndall Watson, North Carolina State Archives
Brent West, University of Illinois
Brian Watts, Peraton