EMAIL ARCHIVES: BUILDING CAPACITY AND COMMUNITY

EMAIL ARCHIVING SYMPOSIUM

June 13-15, 2023

EMAIL ARCHIVING SYMPOSIUM

ABOUT THE SYMPOSIUM

The Email Archiving Symposium is the culminating event of the four-year Email Archives: Building Capacity and Community (EA:BCC) regrant program. This program aimed to enhance email archiving capacity in archives, libraries, and museums by funding projects ranging between 25,000 to 100,000 in work. The symposium provides a unique opportunity to showcase and highlight the EA:BCC funded projects, as well as ongoing work within the global community, fostering collaboration and inspiring future initiatives in the field.

Email archives are a valuable source of information and evidence. More and more institutions are grappling with the challenges of acquiring, arranging, describing, preserving, and providing access to it. Others are just beginning their journey to realize the full potential of email archives as a resource. Age-old questions are being looked at anew. The goal of the EA symposium is to highlight the ways that now, more than ever, we can and must fully integrate email preservation and access into archival practice.

KEYNOTE SPEAKER

NICOLE HEMMER

Nicole Hemmer is director of the Carolyn T. and Robert M. Rogers Center for the American Presidency and associate professor of history at Vanderbilt University.

She is author of *Messengers of the Right: Conservative Media and the Transformation of American Politics* and *Partisan: The Conservative Revolutionaries Who Remade American Politics in the 1990s*. She co-hosts the podcasts Past Present and This Day in Esoteric Political History and is a columnist at CNN.

SYMPOSIUM AT A GLANCE

DAY ONE	Tuesday, June 13, 9:00 – 4:45 PM (CT)

9:00 – 9:30 — **Welcome**
Chris Prom and Ruby Lorraine Martinez

9:30 – 10:30 — **Keynote Address: Beyond the Paper Trail: Doing History in Digital Archives**
Nicole Hemmer

10:30 – 11:00 — **Social**
Facilitated by: Eden Irwin, Ruby Martinez, Chris Prom

11:00 – 11:45 — **ePADD Plus: Integrating Preservation Functionality into ePADD**
Tricia Patterson, Stephen Abrams, Sally DeBauche, Paul Carlyle – EA:BCC Project Briefing

12:00 – 1:00 — **Lunch/Break**

1:00 – 1:45 — **Be Prepared: CoSA's Email Processing Capacity Building**
Nick Connizzo – EA:BCC Project Briefing

2:00 – 2:45 — **Hands-On Email Preservation with Mailbag**   **Using ePADD to Process Email Archives**
3:00 – 3:45 — *Gregory Wiedeman – Workshop (2:00-3:30)*   *Sally DeBauche – Workshop (2:00 – 5:00)*

4:00 – 4:45 — **Email and Institutional Archives**
Andrew Waugh – Project Briefing

DAY TWO	Wednesday, June 14, 8:00 – 4:45 PM (CT)

8:00 – 8:45 — **Email in French, Nlp and Pêlé-mél project**
Grailles Bénédicte, Aït El Mekki Touria – Project Briefing

9:00 – 9:45 — **Email as a Record?!: building institutional support for email archiving at a small liberal arts college**
Jessika Drmacich – Lightning Talk

Why should we archive government email as correspondence? An analysis of the Capstone approach in the light of German historical source studies
Maria Benauer – Lightning Talk

10:00 – 10:45 — **EMCODIST: Introducing the 'Digitally Curious' to Email Archives for Organizational Research and History**
David Kirsch – EA:BCC Project Briefing

11:00 – 11:45 — **PREPAREing Maine for the Future**
Nick Connizzo, Lisa Erickson-Harris – Project Briefing

12:00 – 1:00 — **Lunch Break**

1:00 – 1:45	**Providing Access to Email Archives via Remote Emulation** *Ethan Gates - Lightning Talk*
	Hybrid Moments: Letters and Email in Conversation *Kathryn Antonelli - Lightning Talk*
2:00 – 2:45	**Creating Email Archives from PDFs - The Covid-19 Corpus** *Benjamin Lis – EA:BCC Project Briefing*
3:00 – 3:45	**Preserving Email Attachments with Attachment Converter** *Matt Teichman, Nathan Mull – EA:BCC Project Briefing*
4:00 – 4:45	**Ask an Email Archivist** *Nick Connizzo, Greg Wideman, David Kirsch*

DAY THREE   *Thursday, June 15, 8:00 – 4:00 PM (CT)*

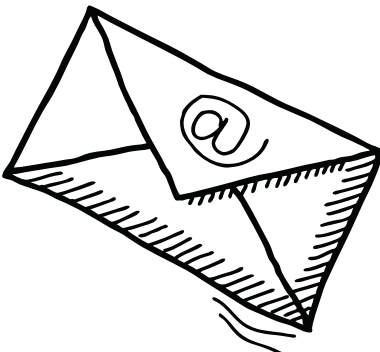| 8:00 – 8:45 | **Archiving mailboxes at the Archives nationales (France): overview and perspectives**
Emeline Levasseur, André Falut – Project Briefing |
9:00 – 9:45	**Developing Email Preservation Training Resources for the Community: the new Novice to Know-How from the Digital Preservation Coalition** *Sharon McMeekin – Project Briefing*
10:00 – 10:45	**The Palladium Project – appraisal, access and use of Carcanet Press email archive** *Jessica Smith, Jochen Farwer – Lightning Talk*
	Email Handling in a Digital Preservation System *Jack O'Sullivan – Lightning Talk*
11:00 – 11:45	**- PDF-M Briefing** *Tom Habing, Chris Prom – Project Briefing*
12:00 – 1:00	**Lunch Break**
1:00 – 1:45	**Time for a Change: Exploring New Avenues for Email Collection Preservation** *Lynda Schmitz Fuhrig - Lightning Talk (Pre-Recorded)*
	Understanding the visualisation needs of researchers using email archives *Peter Green - Lightning Talk (Pre-Recorded)*
2:00 – 2:45	**"Good Enough" is Pretty Darn Good - Working With What You've Got to Move Email Processing Forward** *Laurel McPhee, Tori Maches - Project Briefing*
3:00 – 3:45	**Love the Words: Preserving the Email Collection of 92Y's Unterberg Poetry Center** *Sally DeBauche, Marian Clarke – EA:BCC Project Briefing*
3:45 – 4:00	**Closing Remarks** *Chris Prom*

DETAILED PROGRAM

SYMPOSIUM DAY 1 - JUNE 13, 2023

9:00 AM - 9:30 AM (CT)

Opening Remarks
Chris Prom and Ruby Martinez

9:30 AM - 10:30 AM (CT)

Keynote Address – Beyond the Paper Trail: Doing History in Digital Archives
Nicole Hemmer

As historians begin tackling the work of analyzing the 1990s and early 21st century, they are learning that the traditional paper trail – the memos, records, and correspondences that make up so much of the historian's craft – peters out, and that a world of born-digital sources takes its place. Few historians have been trained to tackle these new archives; even locating emails, videos, and websites remains a monumental task. Nicole Hemmer, a historian of the near-past, will talk about her experiences working with born-digital sources and the challenges the profession must overcome as it reaches the end of the paper-trail era.

10:30 AM - 11:00 AM (CT)

Social
Led by: Eden Irwin, Ruby Martinez, and Chris Prom

Join us for a social to interact with other attendees and the symposium organizers! We will engage in a friendly competition to see who has the smallest AND largest email inbox contest, to share some email war stories, and to look ahead to the rest of the symposium.

1. *Smallest and Largest Email Inbox Contest: Think you have the tidiest and most streamlined email inbox? Or perhaps your inbox is a digital treasure trove overflowing with valuable information? Engage in our friendly competition to determine who boasts the smallest and largest email inbox among the attendees. Showcase your organizational skills, share your strategies, and let's see who comes out on top!*

2. *Email War Story Exchange: Every email professional has experienced moments of triumph, frustration, and humor while managing their inbox. This is your chance to gather with like-minded individuals and share your most compelling email war stories. Whether it's a hilarious misunderstanding, an extraordinary email retrieval, or a near-disastrous mishap averted, we want to hear your tales.*

11:00 AM - 11:45 PM (CT)

ePADD Plus: Integrating Preservation Functionality into ePADD
Tricia Patterson, Stephen Abrams, Sally DeBauche, Paul Carlyle

This briefing will cover the objectives and outcomes of the EA:BCC-funded project "Integrating Preservation Functionality into ePADD" (or ePADD+ project, for short). From 2021-2022, Harvard University, the University of Manchester, and Stanford University collaborated to enhance the open-source email archiving tool, ePADD, with significant new preservation features, including import and export of a wider array of email formats for preservation purposes, capture of complete original order, processing documentation via PREMIS metadata, and Bagging for repository deposit.

The session will commence with an overview of how the collaboration coalesced, as well as the purpose and process of how the project unfolded. Presenters will review the new functions and features available in ePADD version 10.0, including a new optional integration of a format conversion tool that enables the ingest and export of a wide variety of email formats. The session will include a walkthrough of one institutional implementation workflow of ePADD, as well as an overview of how to get started using ePADD at your own organization. Presenters will conclude with a look ahead at next steps in tool and community development and guidance for how others can get involved.

This session is useful for curators, archivists, and digital preservationists interested in utilizing the ePADD tool to create preservation packages for deposit into a digital repository, as well as for any email archiving community members interested in keeping abreast of developments in ePADD.

12:00 PM - 1:00 PM (CT)

Lunch/Break

1:00 PM - 1:45 PM (CT)

Be Prepared: CoSA's Email Processing Capacity Building
Nick Connizzo

Managing and preserving email in government has been a herculean task that 50 states, 6 territories, and the District of Columbia have been dealing with for decades. Questions about proper use of email, separation of personal and public information, handling of protected and sensitive information, and access methodologies continue to challenge government records managers and archivists. PREPARE sought to address these issues of technology and governance with a three-phase project: assessing needs, testing tools, and providing direct assistance in the form of mentoring and best practices related to email preservation and management throughout the enterprise of government.

2:00 PM - 3:30 PM (CT)

Hands-On Email Preservation with Mailbag
Gregory Wiedeman

Learn how to preserve email with Mailbag! This workshop will walk participants though installing and running the mailbagit tool to package email into a mailbag while automatically generating derivatives formats that are compatible with existing tools and workflows. Participants are encouraged to bring their own email exports to try and sample datasets will also be available.

2:00 PM - 5:00 PM (CT)

Using ePADD to Process Email Archives
Sally DeBauche

This workshop will introduce and build familiarity with the ePADD software and provide participants with the experience and tools to advocate for and implement ePADD at their own institutions to appraise, process, and support discovery and access for email archives. The majority of the workshop will be devoted to demonstration and hands-on exercises taking users through specific ePADD modules and functionalities. The workshop will be beneficial for all those with a responsibility or interest in appraising, processing and/or supporting discovery and access for email archives.

4:00 PM - 4:45 PM (CT)

Email and institutional archives
Andrew Waugh

This presentation will describe the results of email appraisal experiments the Public Record Office Victoria ran as an archive tasked with preserving government records. The intended audience is institutional repositories and tool developers.

We will cover:

* The difference between preserving email of individuals and preserving institutional email. We highlight that the concept of email original order is complex, and focussing on the email collections of individual institutional recipients and the way they arranged email may not be optimal.

* The difference between 'positive appraisal' (selecting important emails to preserve) and 'negative appraisal' (selecting the junk to be disposed of). We tie this distinction to the important ability to generalise appraisal tools and approaches (especially AI systems) between different agencies generating email and different jurisdictions.

* The usefulness of email threading as an appraisal tool. Threading reduces the number of appraisal decisions, increases the amount of information on which to base decisions, and provides a better outcome for researchers. We introduce the concept of 'super-threading' as a potential AI area.

* The value of eDiscovery (and similar) tools in processing email collections.

* The difficulty in developing data sets for testing email processing approaches. Privacy and sensitivity concerns have to be balanced against the representative nature of the test collections. This has important consequences for the generalisability of the approaches being developed.

We conclude that email appraisal is a wicked problem to do at scale, especially if you wish to share tools.

SYMPOSIUM DAY 2 - JUNE 14, 2023

8:00 AM - 8:45 AM (CT)

Email in French, Nlp and Pêlé-mél project
Grailles Bénédicte, Aït El Mekki Touria

The Pêle-Mél project, supported by the French Ministry of Culture and led by the University of Angers, has tested strategies for exploring corpora of French emails in order to promote their evaluation and exploitation.

The Pêle-Mél project addressed the issue of electronic mailboxes from a resolutely interdisciplinary perspective, combining expertise in artificial intelligence (Leria laboratory) and archival science (Temos laboratory), and in collaboration with the archives mission of the Ministry of Health and the École des Chartes.

Pêle-mél's objective is to help the archivist to interpret and contextualize an individual mailbox or several mailboxes by categorizing and classifying the mails. To do this, it has developed a method that involves the extraction of named entities, mainly individuals and companies, the extraction of terms, the classification of terms and messages via word embeddings and document embeddings, i.e. by mobilizing a number of tools and techniques of natural language processing. The method is based on supervised machine learning and uses neural network models. It takes into account the linguistic specificities of French.

The program has developed two prototypes, one to classify, the other to query metadata and explore content. Although the experimentation was based on a specific field – the Ministry of Health – the results are generalizable to other environments.

9:00 AM - 9:45 AM (CT)

Email as a Record?!: building institutional support for email archiving at a small liberal arts college
Jessika Drmacich – Lightning Talk

Due to the modern deluge of information and the sometimes ephemeral nature of email as a container, email and other data is often sidelined as a copy of record at small liberal arts colleges who aren't subject to the same record policies of public universities.

Regardless, email is a critical form of communication, as higher education institutions seek to document, preserve, and make accessible their diverse institutional histories. Although not always a record, email exposes information on what is shared and received, as well as how it was communicated.

Williams College is clear in its records policy and retention schedules that some emails and other born-digital materials are records of the college. Yet, compliance can be complicated. Jessika Drmacich, records manager and digital resources archivist at Williams College, will present a toolkit and workflows showcasing how to build support for email archiving and share a low-barrier way to collect and preserve emails.

Why should we archive government email as correspondence? An analysis of the Capstone approach in the light of German historical source studies
Maria Benauer – Lightning Talk

Email has become a key instrument of modern administration and can hold an archival value by providing evidence on government transactions. However, email was not intentionally designed as an administrative mean but rather is a multifunctional tool that can be used to exchange diverse types of information for different purposes. In recent years, multiple international studies reported that this versatile nature of email combined with an increase in quantity has led to a deterioration of recordness in public administrations across the world. In reaction to this, several public archives have implemented the so-called Capstone Approach as a new, automated approach to manage and capture emails of archival value for preservation.

Based on a triangulation of international archival theory and an Austrian source studies approach, I want to examine the Capstone approach's implications for government email as a historical source genre. The Capstone approach does not rely on the end user assigning email messages to single business processes, but shifts the focus to their functional context by appraising email accounts. It thus blurs the line between private and official recordkeeping principles and practices by suggesting structural similarities between government email and personal correspondence. Overall, these findings demonstrate that email archiving is more than a technical challenge. They urge practitioners to reflect on the impact their appraisal decisions may have on the scope of use, interpretability, and evidential function of email archives; and draw attention to practical and conceptual consequences for historical scholarship.

10:00 AM - 10:45 AM (CT)

EMCODIST: Introducing the 'Digitally Curious' to Email Archives for Organizational Research and History
David Kirsch

This project focuses on how users -- of various types, but especially what Talboom & Underdown (2019) call the "digitally curious" -- discover and interact with responsive content in large-scale, organizational email corpora. We have tested our prototype tool EMCODIST (E-Mail COntext DIScovery Tool) with several groups of users, including an IRB approved observational environment where research subjects (users) are asked

to compare their search experience under different information conditions. In this session, we will review the design and implementation of EMCODIST, discuss progress and challenges we have encountered, outline additional steps to be completed by the end of the calendar year, and share high-level findings arising from the work. PI David Kirsch from the University of Maryland, College Park, will present and be joined by co-investigator Adam Nix from the University of Birmingham (UK).

11:00 AM - 11:45 PM (CT)

PREPAREing Maine for the Future

Nick Connizzo, Lisa Erickson-Harris

The PREPARE Project of the Council of State Archivists included community building activities focused on providing direct assistance to states and territories in building digital preservation capacity. We will provide an overview of one such project (designing an email preservation workflow for the State of Maine) including current state, goals, and how the state plans to achieve those goals with the lessons learned from the PREPARE Grant.

12:00 PM - 1:00 PM (CT)

Lunch/Break

1:00 PM - 1:45 PM (CT)

Providing Access to Email Archives via Remote Emulation

Ethan Gates - Lightning Talk

The grant-funded EaaSI (Emulation-as-a-Service Infrastructure) program of work, hosted by Yale University Library, has been working for the past 5 years to build an open-source access platform for remote access to legacy software and data in emulation. Since the underlying technology of the platform can be used to authentically render and interact with a huge variety of domain-specific applications and collections, the team frequently uses case studies to gather feedback and drive development, making sure feature and workflow design meets the real-life needs of archivists, curators, and researchers.

This lightning talk will focus specifically on EaaSI's investigations into using emulation for access to email archives contained within archival and special collections. A recent Yale patron request that included access to a prominent legal scholar's mid-2000s Outlook data, sparked discussion between special collections staff and the EaaSI team regarding both the benefits and limitations of using the EaaSI platform as a type of "virtual reading room". We will cover the current state of the platform and the degree to which it met both staff and patron's needs in this use case, as well as planned development goals the EaaSI team has drafted to better address working with email archives in the future.

I hope this presentation will simultaneously inspire attendees to consider/evaluate how emulation could become a practical part of their own email assessment and access workflows, and for the community around this symposium to dream big together: what would the ideal access platforms of the future look like?

Hybrid Moments: Letters and Email in Conversation
Kathryn Antonelli - Lightning Talk

The recent acquisition of the Richard Garwin Papers by the American Philosophical Society has brought to light the possibilities of more intentionally integrating the analog and digital portions of a 20th century-spanning collection. This lightning talk will examine the role email plays within this hybrid collection. The talk will mainly be targeted toward fellow users and/or potential adopters of ePADD, and will touch upon its correspondent browsing/editing, entity analysis, and authority control capabilities. It will also be of interest to any archivist struggling to straddle the analog-digital divide. The session will ultimately demonstrate how email workflows can be integrated with traditional paper processing to present a finding aid that acknowledges how correspondence can change while correspondents remain the same.

2:00 PM - 2:45 PM (CT)

Creating Email Archives from PDFs - The Covid-19 Corpus
Benjamin Lis, Columbia University

This talk will discuss the technical aspects of an EABCC-funded project to build a Python library that extracts individual email metadata and text from PDFs. We will also report our experiences using it to enhance a corpus of FOIAed documents, some thousands of pages long, from the early days of the COVID-19 pandemic.

3:00 PM - 3:45 PM (CT)

Preserving Email Attachments with Attachment Converter
Matt Teichman, Nathan Mull

Opening old email in whatever software formats are available today is easy enough on the computers we currently have. But what happens when the technology moves on? Will we still be able to read the word processing documents, spreadsheets, PDFs, or graphics files of the past?

Even today, there are many file formats that are difficult to open on modern machines. Because the technology evolves so rapidly, in certain cases it can be challenging even to open a file that is only 10 or 20 years old, to say nothing of 40 or 50.

Thanks to the financial support of the EABCC, we are getting ready to release an open-source tool called Attachment Converter, which batch-converts the attachments in your email collection to digital formats that last longer. Attachment Converter lets you use the file

format conversion utilities that are already installed on your system to

create a copy of your entire email collection, featuring converted copies of every attachment you are able to convert, in the emails from which they originated.

4:00 PM - 4:45 PM (CT)

Ask an Email Archivist

Nick Connizzo, Gregory Wiedeman

In this session, attendees will have the opportunity to ask questions to a panel of EA:BCC awardees. This is an informal setting where attendees can engage with speakers from the symposium and discuss various topics related to email archiving. Our panel of speakers include professionals from different organizations who have successfully completed email archiving programs, as well as researchers and academics who are studying this area. Attendees are encouraged to come prepared with their questions and engage in an open dialogue. To submit your questions beforehand, please submit them here.

SYMPOSIUM DAY 3 - JUNE 15, 2023

8:00 AM - 8:45 AM (CT)

Archiving mailboxes at the Archives nationales (France): overview and perspectives

Emeline Levasseur, André Falut

The Archives nationales (France) have been collecting mailboxes in PST and Mbox format since 2017, and have ingested 40 of them in their digital repository, based on the Vitam software and the SEDA standard, since its opening in 2018. In order to do so, they came up with a mailbox-specific preservation strategy.

This strategy involves transferring each mailbox to the repository as a set of two SIP (prepared by the archival services of ministries, or by the Archives nationales for data migration), one containing a raw export, the other containing the individual messages in both EML and TXT format, as well as their extracted metadata and attachments. The latter SIP is obtained by processing the mailbox in the SIP preparation tool ReSIP, developed by the Vitam team.

The goals are to ensure the preservation of messages and their attachments and to make it possible to answer access requests. As the first access requests on mailboxes recently came in, current reflection is particularly focused on this aspect.

While effective, this strategy has limitations and is meant to evolve as the Archives nationales research means of access to archived mailboxes, monitor relevant tools and publications and experiment.

This intervention is intended for an audience of archivists involved in collecting, preserving and giving access to mailboxes or interested in that topic. We envision it as a talk in three parts: a summary of the Archives nationales' institutional and technical context, a presentation of the preservation strategy and an overview of evolution prospects.

9:00 AM - 9:45 AM (CT)

Developing Email Preservation Training Resources for the Community: the new Novice to Know-How from the Digital Preservation Coalition
Sharon McMeekin

Between December 2019 and March 2021, the DPC developed a new digital preservation training learning pathway on behalf of The National Archives (UK): "Novice to Know-How: Digital Preservation Skills for Beginners". While the course was developed with the UK Archives Sector in mind, it was made available for free to all. It has proved more successful than we could ever have hoped. It has now been completed by nearly 3000 learners in more than 60 countries. The full content of the learning pathway has also been shared with 9 other organizations for inclusion within their own learning management systems. Feedback has shown the course to be accessible, practical, and confidence building.

In late 2022, The National Archives (UK) awarded the DPC a further contract to develop a follow-on learning pathway, focusing on email preservation. The new learning pathway aims to provide learners with all the skills they will need to establish an email preservation program at their organization. It includes content covering organizational issues as well as developing workflows and demos of key tools. As with the original learning pathway, resources will be made available under a Creative Commons license for reuse by the community. These will include illustrations, resource lists, tool demo videos, and SCORM packages of learning content.

In this presentation we will discuss the motivations for the creation of the learning pathway and the development process. We will also provide an overview of the resources available to community and how they can be accessed.

10:00 AM - 10:45 AM (CT)

The Palladium Project – appraisal, access and use of Carcanet Press email archive (Lightning Talk)
Jessica Smith, Jochen Farwer

This proposal is submitted to present on the Arts Council England funded Palladium Project carried out at the University of Manchester in 2021/22.

The University of Manchester Library has been preserving the archive of indie poetry publisher Carcanet Press since 2012. The collection is held in our digital preservation system, Preservica, and 2021 saw the introduction of appraisal and management via ePADD. The Palladium project was initiated to enable the creative arts and project archivists to explore the challenges of access within the context of the regulatory framework, and to explore new software available for the management of email archives.

The presentation will showcase the project's findings relating to appraisal of email, methods of providing mediated access, and collaboration with donors, researchers, artists and developers.

It will outline the development of new functionality for ePADD by the Library's development team, including display and download functionality, email text redaction and correspondent editing. Short demonstration videos of some of these developments will be included.

It will give an overview of the project outputs, including new workflows for preservation and appraisal, an appraisal rubric for the Carcanet email archive, user guides for donors and archive professionals, data visualisations and creative responses to the email archive.

Email Handling in a Digital Preservation System (Lightning Talk)
Jack O'Sullivan

In this talk, we will discuss how Preservica, a COTS Digital Preservation system, handles email. We will look at how we model the information content of emails and how that enables us to perform migration and rendering. We will also discuss how we manage bulk mailbox transfers from leading providers, as well as individual email messages.

11:00 AM - 11:45 PM (CT)

PDF/Mail – Format Overview, Tool Architecture, and Future Directions
Tom Habing, Chris Prom

The EA-PDF (Email Archiving with PDF) project is developing a PDF specification (PDF/mail) for email archiving, as well as an open-source tool to convert emails to the new PDF format. The project aims for PDF/mail to lower barriers to effective email preservation that meets the needs of the archive and digital preservation community. The talk will provide a brief technical overview of the PDF/mail specification itself, including a description of the technical architecture and operation of our mailbox to PDF/mail conversion tool, and possible future directions for the tool. With an open-source tool available for converting emails, archivists and other professionals— particularly those working in context where they do not have access to technologies supporting email preservation--will have a straightforward and cost-effective way of preserving emails for posterity, complementing other preservation methods and tools.

12:00 PM - 1:00 PM (CT)

Lunch/Break

1:00 PM - 1:45 PM (CT)

Time for a Change: Exploring New Avenues for Email Collection Preservation
Lynda Schmitz Fuhrig - Lightning Talk (Pre-Recorded)

The Smithsonian Libraries and Archives (SLA) has been collecting and preserving email for more than 15 years and led an early email preservation project called CERP that resulted in an XML preservation format and software applications. Even though tools like Teams and Slack are replacing some email usage, the reliance on email in many organizations has continued for communication. Market research group

Radicati noted that business and consumer email messages per day would be more than 333 billion in 2022 and will reach more than 392 billion in 2026.

The last few years have resulted in various email collection projects that include ePADD+, Mailbag, and RATOM, and literature from the field has focused on topics including email preservation and specifications. This is the right time for SLA to rethink our email collection processes from ingest to processing/appraisal to preservation and access.

SLA has made adjustments in recent years to its workflows to assist in appraisal and accessioning, but backlogs have continued as more frequent and larger collections are acquired thanks to larger mailbox size limits and cloud-based applications. As an early tester of Phase 1 of ePADD, SLA has closely watched developments and was delighted to see the ePADD+ project offering additional features. Participating as a 2022 ePADD+ community tester reinforced the idea of exploring a change.

This lightning talk will focus on the initial results of exploring tools like ePADD+ and incorporating revised workflows. The intended audience is anyone who acquires or is planning the acquisition of email collections.

Understanding the visualisation needs of researchers using email archives
Peter Green - Lightning Talk (Pre-Recorded)

Providing access to email archives for researchers will include providing visualisations which have long been seen as a potential solution to the challenges of working with email archives. However, there is little research that covers the visualisation needs of researchers in this context. The research I will be undertaking as part of my Doctoral thesis will help to close that knowledge gap. I will identify a cohort of researchers, experienced in working with correspondence held in cultural institutions, and engage with them using semi-structured interviews within a qualitative research methodology. These interviews will explore how the researchers go about their research, what they need to help them achieve their research goals, and what forms of email visualisation would best meet their needs. The outcomes of this research will be used to guide the development and deployment of visualisation of email archives in cultural institutions and will fill a significant knowledge gap.

In the Lightning Talk I will present a brief overview of my proposed research and timeline. I hope to get questions and feedback, perhaps offline after the session, that will help inform my research.

This research will be of interest to anyone who is managing email archives and pondering how they will be used by researchers in the future.

2:00 PM - 2:45 PM (CT)

"Good Enough" is Pretty Darn Good - Working With What You've Got to Move Email Processing Forward
Laurel McPhee, Tori Maches - Project Briefing

Processing and providing access to email continues to be a major bottleneck in many born-digital archives programs. With the logistical and technical challenges involved, it can be easy to lose sight of simpler strategies to make email available to

researchers now. Are you stuck feeling like you're waiting indefinitely for a tool or system that provides a perfect "email experience?" This briefing will help you think about other ways to move forward.

This session highlights steps taken by archivists at UC San Diego to process and provide "good-enough" access to email across a variety of format types, utilizing the Mailbagit open source tool and the DAMS system already in place at the library. They will describe the process of building confidence with email via working with individual messages saved as plaintext or images, and using this to develop a format-agnostic approach to email in collections.

Presenters will provide a case study of building on this foundation to process an email collection from the 1990s. We'll show how we create PDF surrogates to facilitate keyword searching, keep threads intact, and leverage existing infrastructure and resources to facilitate discovery and remote access in the short term, while preserving the original files.

3:00 PM - 3:45 PM (CT)

Love the Words: Preserving the Email Collection of 92Y's Unterberg Poetry Center

Sally DeBauche, Marian Clarke – Project Briefing

A current count puts the 92nd Street Y's Unterberg Poetry Center's email archive at almost 3 million messages. The email collection offers a look into the day-to-day activities and program planning of what many consider to be the preeminent place for writers and poets to share their work with the public. A collaboration between the Poetry Center and the digital archives team at Stanford University Libraries, this project applies ePADD to the assessment and preservation of the email archive with the goal of developing a processing and accessibility model that other cultural centers might learn from and adopt for their own internal and external purposes. The project team will provide an update on our progress to date, the various challenges that have arisen in a project of this nature, and the creative solutions that we have implemented. We will discuss the process of acquiring the 92Y's email accounts, securely transferring them between project partners, and ingesting them into ePADD. We will also share our approach to processing this large and complex collection to make it usable for stakeholders at the 92Y and beyond, as well as the potential scholarly value of the collection as a record of the public history of literature.

4:00 PM - 4:15 PM (CT)

Closing Remarks
Chris Prom