The Future of Email Archives: A Report from the Task Force on Technical Approaches for Email Archives

Email archives, like most born-digital records, have added challenges to the traditional practices that many archivists have relied on to manage collections and make them available for research. The inherent properties of email, its variatious platforms, and differences in usage often cause unstructured, reactive efforts to manage email archives as archivists are additionally confronted by time, money, storage, personnel, and training constraints. Without veiling complexities, *The Future of Email Archives* instead offers solutions to acknowledged challenges while simultaneously persuading its readers to be proactive participants in the collection management of email archives.

In 2016, the Andrew W. Mellon Foundation and the Digital Preservation Coalition (DPC) announced the formation of the Task Force on Technical Approaches for Email Archives[1] to examine current efforts to preserve email, articulate frameworks for email preservation, and develop a working agenda to construct the technical framework. Cochaired by Christopher Prom (University of Illinois at Urbana-Champaign) and Kate Murray (Library of Congress), the task force includes a total of nineteen archivists, librarians, technology specialists, and historians from academic, government, technology, and museum sectors. The size of the task force and the diversity of its members enable a successful blend of varied perspectives and areas of expertise into a seamless reference piece helping answer the question: what should archivists do with email?

Recent literature, case studies, and informal writings, such as blog posts and institutional reports, have spiked in the last few years revealing emerging trends in the professionalization of born-digital archival practice, but few evaluate email quite as holistically as *The Future of Email Archives*. Comparatively, the DPC, as part sponsor of this task force, certainly has an interest in email archives, as its 2011 *Technology Watch Report*,[2] also by Christopher Prom, acts as one of the first instructive pieces to specify how to digitally preserve email. Acting as more of a formal article, it includes a brief literature review and more specific details on email preservation technologies. Yet, in the technology world, a report more than five years old often hints at obsolescence,

and publishing an unofficial sequel like *The Future of Email Archives* fills needed research gaps. Pairing information from the 2011 report with the task force's 2018 report, published by the Council on Library and Information Resources (CLIR), may give archivists a broader background in digital preservation research surrounding email, but it is not necessary to comprehend *The Future of Email Archives*. Furthermore, if readers seek more assistance after reading the report, the task force includes supplementary documentation on its website,[3] defining email tools and standards.

After a year of research, the task force's efforts culminated in this concise, yet comprehensive, report, which critically analyzes the entire email life cycle. In its first draft, the report was designed to answer the why, when, who, where, what, and how of email archives based on five different working groups of the task force.[4] The veins of this structure are still visible in the final report, as it begins by contextualizing why email archives are important and then walks readers through email's technical properties and how it can be managed. While the report itself addresses these questions, it simultaneously calls for a culture shift to the community at large and advocates for the need to invest resources and expertise in this field.

Part 1 of the report, "The Untapped Potential of Email Archives," details the intrinsic importance of emails for the archival record. The most persuasive portion of the report, it admits that its agenda includes a "call to arms" (p. 5) that openly advocates stakeholders' more active participation and future advancements in email preservation. In four brief subsections, the authors emphasize the various risks of losing these messages to time and identify areas of opportunities to preserve them instead. After setting the stage with their thesis upfront, the task force transitions to the rest of the report, which includes broad technical details about email and how it relates to archival theory. Part 2, "The Email Stewardship Lifecycle," introduces the concept of email as a record for both institutional and personal collections in the archival context. Breaking down the details of the general records life cycle, the task force explains how creation/use, appraisal, acquisition, arrangement and description, preservation, and access activities impact email archives.

Transitioning to more details about email itself, Part 3, "Email as a Documentary Technology," elaborates the intricacies of email technologies. It includes an overview of email architecture, its storage properties and design, methods of transmitting messages, security vulnerabilities, and the nature of attached and linked content in messages. This portion is peppered with references to how each of these technologies directly impacts the archival record. Similarly, Part 4, "Current Services and Trends," describes current email services and user trends to promote continuous monitoring and assessment within the

IT industry and society at large. This section explains how the technologies mentioned in Part 3 evolve based on user demands, needs, and expectations. It continues the theme of maintaining archival contextualization as it details how each evolving technology and user trend impacts or may impact archival processes for better or worse. Additionally, Part 4 identifies more specific challenges for repositories in their efforts to collect, authenticate, track administrative actions on the data, ensure security and privacy, and process large collections of email.

After reviewing the technological aspects of email archives in the first two-thirds of the report, Part 5, "Potential Solutions and Sample Workflows," delivers practical, potential strategies for preserving and managing email archives. Broken into three subsections, the first summarizes digital preservation techniques for email including bit-level preservation, migration, and emulation options. Because any preservation activity depends upon the interoperability of various tools and software, the second subsection explains common interoperability features among email and archiving platforms. Finally, it suggests possibilities for implementing workflows for its three aforementioned digital preservation strategies based on current interoperability designs and functionality limitations. While noting that these potential workflows and solutions are feasible for more well-resourced institutions, the task force acknowledges that most institutions are left behind. As a result, Part 6, "The Path Forward: Recommendations and Next Steps," explores and recommends next steps for community-driven research, advocacy, and tool development so that the practice of email archiving can become more ubiquitous and accessible. It identifies specific short- and long-term actions the archival community can take to develop advocacy and tools for email archives. Finally, the report concludes with several appendixes that index recent tools for archiving email and current research projects dedicated to email archives, further emphasizing its message in Part 6 to expand knowledge and research of email archiving efforts.

It was hard not to agree with the task force's arguments that email archives are imperative and that we collectively need to be more proactive about their preservation. As a practicing digital archivist, I have dealt with the apprehension of handling emails from curators, collection managers, and donors alike. But as a traditionally trained archivist, I instead assert the intrinsic importance of correspondence and urge others not to be discouraged by a medium. It was inspiring to read similar opinions and to be encouraged to act by like-minded professionals; I often found myself thinking: "yes! I completely agree!" Initially, I assumed this would only be an instructive reference resource, but I was delighted to see advocacy as the underscore of the report. Nearly every challenge mentioned for archiving and preserving email was met with opportunities on

how to address those challenges. Simultaneously, the humble admission that these opportunities may not be final answers aligns with most digital preservation activities as we continuously work to improve and finalize solutions and best practices cooperatively.

The report itself declares that it is meant for a wide audience, and it certainly can be adapted for many different needs. Most notably, it is particularly useful to novices with entry-level understanding of email archives and preservation as a handy guide that will introduce email archiving without inundating readers; nearly all jargon is explained clearly and likened to archival theory that helps contextualize its properties. Given the strong representation of advocacy for email archiving and preservation throughout the report, it is undoubtedly useful for those who participate in general outreach and donor relations at their institutions. Although the report does not provide exact templates for communicating the importance of email archiving, its general language can be reconfigured and adapted to various needs. Even seasoned digital archivists will benefit from the report as they can adapt its frameworks to workflows, allowing for consistent implementation among various professionals. Like all CLIR reports, its outlined structure makes it easy to skip to sections most beneficial to one's needs.

As a concise report, it is hard to identify glaring faults when it freely admits need for more research and action throughout the profession; in general, the report does not pretend to be something it is not. As daily work evolves, archivists should look to this report as a foundation for framing institutional policies, workflows, strategies, and goals, but not necessarily as a prescription. Significantly, it is poised to act as a springboard for archivists to formalize their practices, policies, and technologies for archiving email and advancing the field overall.

© **Kelsey O'Connell**
Northwestern University Libraries

NOTES

1 Andrew W. Mellon Foundation and Digital Preservation Coalition, "Mellon Foundation and Digital Preservation Coalition Sponsor Formation of Task Force for Email Archives," November 1, 2016, https://mellon.org/resources/news/articles/mellon-foundation-and-digital-preservation-coalition-sponsor-formation-task-force-email-archives.

2 Christopher J. Prom, "Preserving Email," *DPC Technology Watch Report 11-01* (2011), http://dx.doi.org/10.7207/twr11-01.

3 Task Force on Technical Approaches for Email Archives, "Supplementary Documents," March 2018, http://www.emailarchivestaskforce.org/documents.

4 Task Force on Technical Approaches for Email Archives, "Task Force on Technical Approaches to Email Archives Consultation Draft Report," August 31, 2017, https://goo.gl/52cpvf.