

Archiving Email Symposium

Background Reading

June 2, 2015

Overview / General Information

InterPares 3: General Study 05 – Keeping and Preserving E-mail. Final Report 2009:

http://www.interpares.org/ip3/display_file.cfm?doc=ip3_italy_gs05a_final_report.pdf

- The aim of this study is to investigate the technical aspects relevant to the e-mail creation, capture and maintenance (i.e., records management) and permanent preservation (i.e., archival) processes.

Prom, Chris. Preserving Email. DPC Technology Watch Report 11-01 December 2011

<http://www.dpconline.org/publications/technology-watch-reports>

- Excellent coverage of issues with preserving email.

Technical Perspectives

Bernstein, Daniel. Internet mail message header format

<http://cr.yp.to/immhf.html>

- Describes format of an Internet mail message header

Knight. Gareth. InSPECT Significant Properties Testing Report: Electronic Mail. 2009.

<http://www.significantproperties.org.uk/email-testingreport.html>

- Discussion of important characteristics of email objects that would need to be retained in a preservation workflow.

Library of Congress. *The Signal* blog. Email related blog posts.

- Emerging Collaborations for Accessing and Preserving Email (September 19, 2014):
<http://blogs.loc.gov/digitalpreservation/2014/09/emerging-collaborations-for-accessing-and-preserving-email/>
- The ePADD Team on Processing and Accessing Email Archives (October 20, 2014):
<http://blogs.loc.gov/digitalpreservation/2014/10/the-epadd-team-on-processing-and-accessing-email-archives/>
- The “How” of Email Archiving: More Launching Points for Applied Research (July 18, 2013):
<http://blogs.loc.gov/digitalpreservation/2013/07/the-how-of-email-archiving-more-launching-points-for-applied-research/>
- The “What” of Email Archiving (July 2, 2013):
<http://blogs.loc.gov/digitalpreservation/2013/07/the-what-of-email-archiving/>
- Shaking the Email Format Family Tree (April 4, 2014):
<http://blogs.loc.gov/digitalpreservation/2014/04/shaking-the-email-format-family-tree/>

Library of Congress. Sustainability of Digital Formats Planning for Library of Congress Collections. Email related format descriptions, including links to RFCs and other technical documentation.

- cc:Mail Archive Email Format (CCA):
<http://www.digitalpreservation.gov/formats/fdd/fdd000391.shtml>

- Email (Electronic Mail Format, EML):
<http://www.digitalpreservation.gov/formats/fdd/fdd000388.shtml>
- Internet Message Format (IMF):
<http://www.digitalpreservation.gov/formats/fdd/fdd000393.shtml>
- MBOX Family of Email Formats:
<http://www.digitalpreservation.gov/formats/fdd/fdd000383.shtml>
- Microsoft Compound File Binary File Format, Version 3 (CFB):
<http://www.digitalpreservation.gov/formats/fdd/fdd000380.shtml>
- Microsoft Compound File Binary File Format, Version 4 (CFB):
<http://www.digitalpreservation.gov/formats/fdd/fdd000392.shtml>
- Microsoft Outlook Item (MSG):
<http://www.digitalpreservation.gov/formats/fdd/fdd000379.shtml>
- Microsoft Outlook PST 2003 (Unicode):
<http://www.digitalpreservation.gov/formats/fdd/fdd000378.shtml>
- Microsoft Outlook PST 97-2002 (ANSI):
<http://www.digitalpreservation.gov/formats/fdd/fdd000377.shtml>

North Carolina State Archives. Preservation of Electronic Mail Collaboration Initiative. 2009.
<http://www.history.ncdcr.gov/SHRAB/ar/emailpreservation/reports.htm>

- Collaborative effort to develop XML schema for email records, including PST files.

Smithsonian Institution Archives and Rockefeller Archive Center. Collaborative Electronic Records Project. 2009

Project: <http://siarchives.si.edu/cerp/index.htm>

Email parser: <http://siarchives.si.edu/cerp/parserdownload.htm>

XML schema: http://www.history.ncdcr.gov/SHRAB/ar/emailpreservation/mail-account/mail-account_docs.html

- Overview of the CERP project and deliverables, including XML schema.

Stanford University Libraries. ePADD Project. 2015.

<http://library.stanford.edu/spc/more-about-us/projects-and-initiatives/epadd-project>

- Project to develop open-source tool that will allow repositories and individuals to interact with email archives before and after they have been transferred to a repository.

Wikipedia. Email

<http://en.wikipedia.org/wiki/Email>

- Helpful article which discusses the structures and protocol of email messages.

Archival Perspectives

Baker, Fran, Phil Butler; Ben Green. Carcanet Press Email Preservation Project. JISC. 2012.

<https://www.escholar.manchester.ac.uk/api/datastream?publicationPid=uk-ac-man-scw:165096&datastreamId=FULL-TEXT.PDF> and <http://www.dpconline.org/advocacy/awards/digital-preservation-awards-2014/1277-carcanet-press-email-archive-university-of-manchester>

- Email archive of Carcanet Press, UK poetry publishers archive housed at University of Manchester Library (UML). Basing work on both traditional archival practice and digital preservation standards, project established workflows, mechanisms and documentation for

acquiring, migrating and preserving this sizeable email archive, which (after appraisal) extends to 16 GB and approximately 170,000 emails. 2014 DPC Preservation Award Winner.

Cocciolo, A. (2015, under review). Email as Cultural Heritage Resource: Appraisal Solutions from an Art Museum Context.

http://www.thinkingprojects.org/Email_as_Cultural_Heritage_Resource_forweb.pdf

- This study examines strategies to appraise significant email as a cultural heritage resource, using an art museum as a study site.

Kenneth Hawkins, "A Pattern Language for Electronic Records," *Archivaria* 67 (Spring 2009), 157-188

Library of Virginia's Kaine Email Project

<http://www.virginiamemory.com/collections/kaine/>

- Provides access to email records and how staff made email records available.

Lee, Christopher A., and Kam Woods. Automated Redaction of Private and Personal Data in Collections: Toward Responsible Stewardship of Digital Heritage. In *Proceedings of Memory of the World in the Digital Age: Digitization and Preservation: An International Conference on Permanent Access to Digital Documentary Heritage*, 26-28 September 2012, Vancouver, British Columbia, Canada, edited by Luciana Duranti and Elizabeth Shaffer, 298-313: United Nations Educational, Scientific and Cultural Organization, 2013.

<http://ils.unc.edu/callee/p298-lee.pdf>

- Discusses overarching issues of personal and private data, including email records, and BitCurator's functionality to deal with these data types in an automated process.

National Archives (UK). Redaction toolkit: Editing exempt information from paper and electronic documents prior to release

http://www.nationalarchives.gov.uk/documents/information-management/redaction_toolkit.pdf

- Provides guidance on editing exempt material from information held by public bodies in the UK, including email records.

Records Management Perspective

ARMA International. Best Practices for Managing Electronic Messages (technical report) – fee

<https://members.arma.org/eweb/browse.aspx?site=armastore&webcode=product&id=9f39921b-b13d-41e7-a4db-23e56f42eddb#.VVntKWN7TSg>

- This technical report provides current information for managing all types of text-based electronic messages or communication, such as e-mail, instant messages, and text messages (SMS). Includes a model electronic message management program audit guideline.

InterPares 3: Guidelines and Recommendations for E-Mail Records Management and Long-Term Preservation. Final Report 2011.

http://www.interpares.org/ip3/display_file.cfm?doc=ip3_italy_gs05b_final_report.pdf

- Overview of the records management issues with email as a record along with recommendations on business practices and workflows. Extensive bibliography.

National Archives and Records Administration. Email Management resources.
<http://www.archives.gov/records-mgmt/email-mgmt.html>

- Includes links to Capstone.

Pennock, Maureen. Curating E-Mails: A life-cycle approach to the management and preservation of e-mail message. DCC | Digital Curation Manual. 2006.

<http://www.dcc.ac.uk/sites/default/files/documents/resource/curation-manual/chapters/curating-e-mails/curating-e-mails.pdf>

- Guides for managing and curating e-mail messages for both current and future use, including generic framework for e-mail curation and preservation, provides a summary of current approaches, and addresses the technical, organizational and cultural challenges to successful e-mail management and longer-term curation.

Policy Perspectives

National Council of Nonprofits. Document Retention Policies for Nonprofits

<https://www.councilofnonprofits.org/tools-resources/document-retention-policies-nonprofits>

- Outlines policies for non-profits, including email.

Society of American Archivists. Statement on Use of Non-government Email Accounts for the Conduct of Public Business

<http://www2.archivists.org/news/2015/statement-on-use-of-non-government-email-accounts-for-the-conduct-of-public-business>

- Statement from SAA on the practice of elected and appointed government officials of using unofficial, non-government email accounts for the conduct of public business.