Practical Technology for Archives

Processing Unidirectional E-Mail Memos for Preservation and Access

Lisa M. Schmidt
Michigan State University

Processing Unidirectional E-Mail Memos for Preservation and Access

Description

[Excerpt] As with most institutions, Michigan State University now communicates official university business through e-mail messages rather than traditional paper memos. The university's IT Services department periodically sends out aggregations of these Deans, Directors, and Chairs (DDC) messages that originate in various offices and departments. Many of the messages include supplementary information as attachments, typically in PDF, DOC, XSL, and/or PPT file formats.

Per MSU's retention schedules, DDC messages have historical value and must be preserved and made accessible. The original file formats—MSG (Windows PC) or EML (Apple Macintosh)—could be converted to MBOX or EML (if necessary), de facto preservation format standards for e-mail, and accessed in nearly any e-mail program. With the DDC messages functioning as unidirectional memos, however, the University Archives & Historical Collections (UAHC) decided to take a simpler approach and convert them to PDF format. The original appearance of the message would be retained, including the header.

Keywords

digital archiving, preservation, e-mail memos

Processing Unidirectional E-Mail Memos for Preservation and Access

Lisa M. Schmidt

Michigan State University

As with most institutions, Michigan State University now communicates official university business through e-mail messages rather than traditional paper memos. The university's IT Services department periodically sends out aggregations of these Deans, Directors, and Chairs (DDC) messages that originate in various offices and departments. Many of the messages include supplementary information as attachments, typically in PDF, DOC, XSL, and/or PPT file formats.

Per MSU's retention schedules, DDC messages have historical value and must be preserved and made accessible. The original file formats—MSG (Windows PC) or EML (Apple Macintosh)—could be converted to MBOX or EML (if necessary), de facto preservation format standards for e-mail, and accessed in nearly any e-mail program. With the DDC messages functioning as unidirectional memos, however, the University Archives & Historical Collections (UAHC) decided to take a simpler approach and convert them to PDF format. The original appearance of the message would be retained, including the header.

UAHC chose the PST Viewer software to normalize the messages to PDF format. PST Viewer provides PDF export functionality, detaching and making available any attachments. Because UAHC uses Archivematica processing software to ingest digital files into its repository system, the PDFs would be normalized to PDF/A format for preservation and the original PDF retained as the access copy. Attachments would be

ingested and normalized for preservation and access following their respective format rules.

Every few weeks, the latest DDC messages are deposited in a specially designated e-mailbox. An archivist moves them to a staging area and uses PST Viewer to normalize them to PDF format. For easier arrangement of the resulting files, the PST Viewer profile is set as shown in Figure 1.
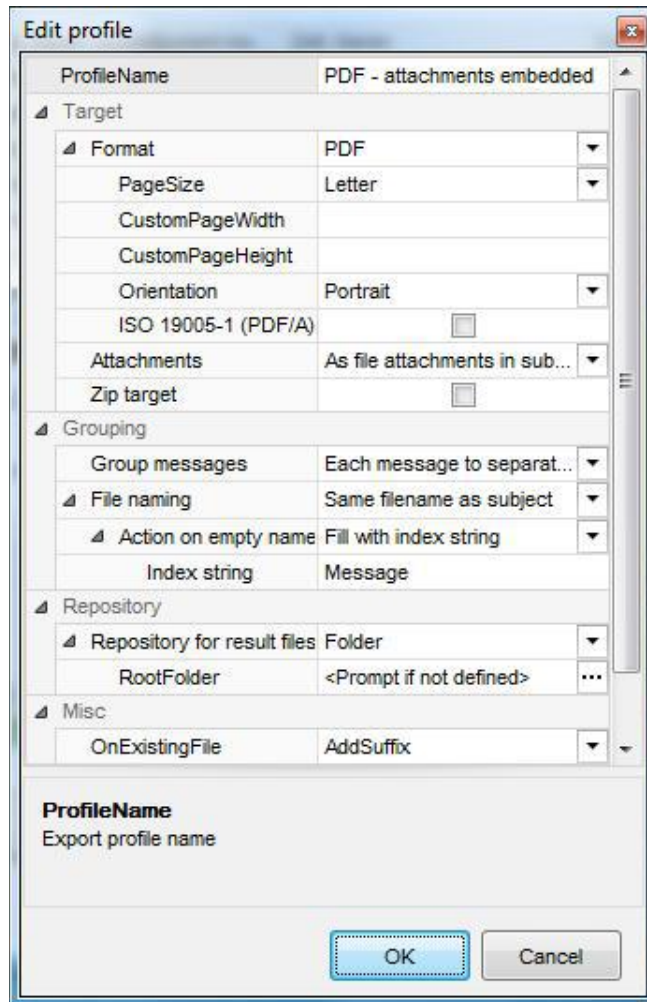

Figure 1. Preferred profile setting in PST Viewer

Within PST Viewer, the archivist navigates to the directory containing messages of interest and exports them to a designated directory. In that directory, messages and attachments are arranged into folders. Names of folders, as well as PDF files of

messages without attachments, are appended with the date of the original message in the yyyy-mm-dd format. Spaces in names are replaced with underscores using the Renamer tool.

Messages are then examined to identify office of origin and moved to folders labeled with that office's record group number, following UAHC's classification system. For example, the record group number for the Office of the President is UA 2.1. Each record group folder is assigned an accession number and ingested into the archival repository as a Submission Information Package (SIP), generating Metadata Encoding and Transmission Standard (METS) files containing metadata; Archival Information Packages (AIPs) containing PDF/As and preservation copies of attachments; and Dissemination Information Packages (DIPs) containing the PDFs and access copies of attachments. Finally, the Archivists' Toolkit collection management system is updated to include the e-mail memos in finding aids.
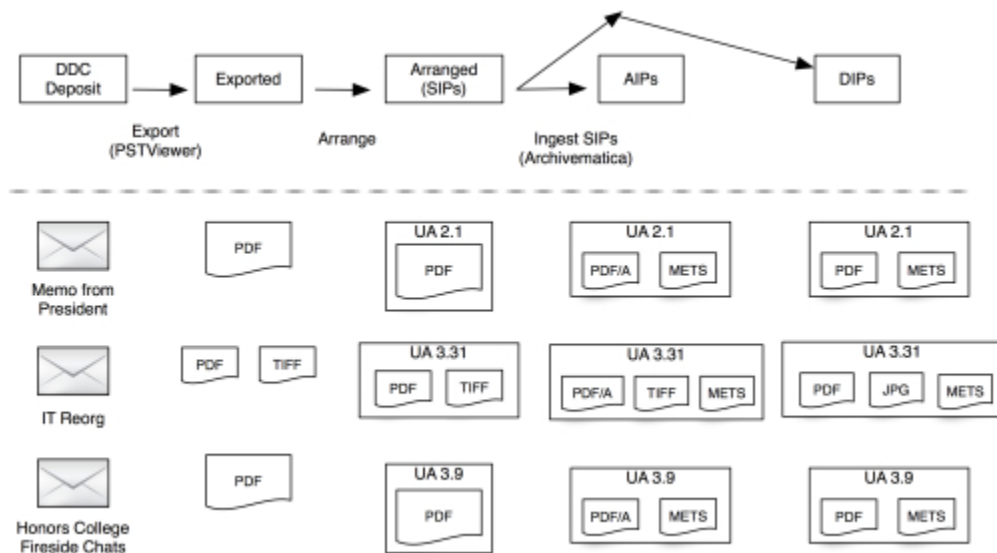

Figure 2. E-Mail Memo Processing Workflow

About the Author

Lisa M. Schmidt is an Electronic Records Archivist at the Michigan State University Archives & Historical Collections. In 2007, she began working for MATRIX, the digital humanities research center at Michigan State, where her work included an NHPRC-funded project to assess and improve the preservation practices for the H-Net academic e-mail lists. She's been at the University Archives since 2010, and her work there includes contributing to the development of a trusted digital repository for electronic institutional records and creating transfer, appraisal, ingest, storage, and access workflows for the university's digital content of enduring value. With MSU librarians, she also provides research data management services to the MSU research community. Lisa holds an MS in Information Studies from the University of Texas at Austin. Prior to joining the information studies field, she enjoyed a successful career as a high technology marketing communications professional in the San Francisco Bay Area and Austin.