

Email Conversion Software Report

Introduction

The McGill University Archives (MUA) is exploring email conversion software in an effort to convert emails from their native and proprietary format to a non-proprietary format in order to ensure long-term preservation and accessibility. This work forms part of the MUA's ongoing *digitalpermanence* initiative.

ABC Amber Outlook Converter¹ allows users to batch convert Outlook/Exchange email folders and messages to other file formats, including proprietary formats such as PDF in addition to non-proprietary file formats such as HTML and ASCII text. ABC Amber Outlook Converter operates in conjunction with Microsoft Outlook/Exchange by reading Microsoft's Outlook Data File (PST files) and mirroring a user's email folder structure in its own interface. Users can choose to convert either email folders or individual email messages to another file format. The output of the conversion process is sent to a user-defined location on a workstation, and the newly converted email messages can be viewed in a web browser (HTML conversion) or other application (PDF conversion).

The MUA tested ABC Amber Outlook Converter (v. 3.06) using PST files from two sources: a McGill professor whose records document teaching/research and administrative functions, and a former McGill University Archives' employee. Our goal was to test the software's ability to convert email messages to HTML, XML, and PDF. In evaluating the software, we examined the output, ensuring the email's header information (i.e. TO, FROM, SUBJECT, DATE) and content (i.e. body of message) were correctly carried over to the new file format and that any email attachments (such as Office-type documents and images) remained accessible and associated with the converted email message. We chose to run our conversion tests on the Sent Items folder on the premise that this was among the fullest accounts of the individual's actions taken, reflected in email replied or sent by the individual. The following is a summary of the preliminary results:

Conversion to HTML

- Product converts email message to individual HTML files/converts all email messages into one large HTML file;
- Email content (i.e. the header and body information) carried to the new file format;
- Attachments slotted to “**_attach**” folder on the same level as the converted email;
- Hyperlinks to attachments created in the new HTML document.

Conversion to XML

- Product converts to XML *but requires a parser or special reader to view the document*;
- Email content (i.e. the header and body information) carried over;
- No hyperlinks were created in the XML document and therefore there is no way for one to click and view an email's attachments.

Conversion to PDF

- Product converts to PDF using a batch processing feature which allows for the conversion of multiple folders;

¹ Available to buy or to try (30 day trial) at <http://www.processtext.com/abcoutlk.html>

- Output in clear presentation style (lending itself to presentation to researcher);
- Attachments slotted to “**_attach**” folder on the same level as the converted email;
- Hyperlinks to attachments were created but do not consistently work.

Conclusion

The product performs conversion functions as promised without a lengthy learning curve to operate. Preliminary testing revealed weakness in ability to maintain hyperlinks to email attachments, particularly in PDF and XML conversions. Concerns were also noted regarding the product's functionality for large-scale batch processing. The MUA ***digitalpermanence*** continues to review possible options for email conversion.

Consultation

The MUA welcomes input from colleagues with a common interest in e-records management.

The following are general resources on email conversion tools:

<http://support.dtsearch.com/faq/dts0169.htm>.

David Kemper

June 2005

www.archives.mcgill.ca