

**Appendix A – Brief overview of tools**

| Tool   | File Format  | Brief technical information   | Attachments?   | Notes  |
|--|--|---|--|--|
| DAVID<br><a href="http://www.expertisecentrumdavid.be/davidproject/teksten/Rapporten/Report4.pdf">http://www.expertisecentrumdavid.be/davidproject/teksten/Rapporten/Report4.pdf</a>                 | XML Schema   |   |  | Flemish project that references European systems             |
| XENA (Xml Electronic Normalising for Archives)<br><a href="http://xena.sourceforge.net/">http://xena.sourceforge.net/</a>  | .pst, .mbox. and .msg converted to ASCII character file in XML wrapper                     | Open-source download available for Windows, Mac or Linux                      |  | Part of a larger product                                     |
| PeDALS (Persistent Digital Archives and Library System)<br><a href="http://www.pedalspreservation.org/Default.aspx">http://www.pedalspreservation.org/Default.aspx</a>                               | .pst files from Microsoft Outlook converted to XML   | Windows (Microsoft Outlook)   | Also converted to XML  | Generated access pdf of each email; part of a larger product |
| CERP (Collaborative Electronic Records Project)<br><a href="http://siarchives.si.edu/cerp/parserdownload.htm">http://siarchives.si.edu/cerp/parserdownload.htm</a>                                   | Convert collection to .mbox (if not already in that format) and then convert to XML Schema | Runs in a virtual machine utilizing Squeak and Smalltalk on PC, Mac, or Linux | Attachments embedded in XML file and exported into a companion directory structure | Handles inactive email, not continually harvested accounts   |
| EMCAP (Email Collection and Preservation)<br><a href="http://www.history.ncdcr.gov/SHRAB/arr/EmailPreservation/default.htm">http://www.history.ncdcr.gov/SHRAB/arr/EmailPreservation/default.htm</a> | Convert collection to .mbox (if not already in that format) and then convert to XML Schema | Windows (requires an MS SQL server)   | Encoded attachments associated with XML file                                       | Handles continually harvested email                          |
| ePADD<br><a href="https://library.stanford.edu/projects/epadd">https://library.stanford.edu/projects/epadd</a>   | Imports directly from accounts or from .mbox files   | Open source software that works on Windows or Mac                             | Able to recognize and display a wide variety of formats                            | Discovery Access portal                                      |

|  |                               |   |   |   |
|--|-------------------------------|---|---|---|
| <p>TOMES (Transforming Online Mail with Embedded Semantics)<br/> <a href="https://www.ncdcr.gov/resources/reco-rds-management/tomes">https://www.ncdcr.gov/resources/reco-rds-management/tomes</a></p>                         | <p>.pst to XML</p>            | <p>Focused on hosted email platforms like Microsoft 365 and Gmail</p> |   | <p>Building on the natural language processing of ePADD</p> |
| <p>EAS (Electronic Archiving System)<br/> <a href="http://library.harvard.edu/lts/systems/eas/features">http://library.harvard.edu/lts/systems/eas/features</a></p>  | <p>Normalized to .eml</p>     |   |   | <p>Tool developed at Harvard for summer 2017 launch</p>     |
| <p>Archivematica<br/> <a href="https://wiki.archivematica.org/Email_preservation">https://wiki.archivematica.org/Email_preservation</a></p>  | <p>.pst to XML or text</p>    | <p>Runs in a virtual machine as part of larger tool</p>               | <p>Converted to preservation and access formats</p> |   |
| <p>Preservica<br/> <a href="http://preservica.com/events/webinars-live-demos/15/07/2015/email-archiving-and-preservation">http://preservica.com/events/webinars-live-demos/15/07/2015/email-archiving-and-preservation</a></p> | <p>.pst and .mbox to .eml</p> | <p>Cloud and Enterprise utilize different processes for ingest</p>    |   |   |