

Appendix

OPF Digital Preservation Community Survey 2019

This reference document is for use with the OPF digital preservation community survey 2019. It is not intended to be an exhaustive list of tools and resources, or provide recommendations. The lists have been extended with each iteration of OPF's member and community surveys to reflect additional tools and resources in use by some organisations.

Q13: OPF Reference Toolset	
JHOVE JHOVE GitHub	An open source file format identification, validation and characterisation tool from the Open Preservation Foundation.
Jpylyzer Jpylyzer GitHub	Jpylyzer is a validator and feature extractor for JP2 images from the Open Preservation Foundation.
FIDO FIDO GitHub	FIDO (Format Identification for Digital Objects) is a command-line tool maintained by OPF and used for identification of digital files based on their PRONOM signature.
veraPDF veraPDF GitHub	veraPDF is a purpose-built, open source, file-format validator covering all PDF/A parts and conformance levels. It is maintained by OPF as part of the open source digital preservation reference toolset.
DPF Manager	A TIFF conformance checker for digital preservation produced by the PREFORMA project.
xcorrSound xcorrSound GitHub	Suite of tools for automated quality assurance of audio migration processes from the SCAPE project.
Q14: Software	
Apache Tika	Format identification / Characterisation / Content extraction
AtoM (Artefactual)	AtoM stands for Access to Memory. It is a web-based, open source application for standards-based archival description and access in a multilingual, multi-repository environment.
Bagit	File packaging and transfer format for arbitrary digital content.
BitCurator	Digital forensic tools for collecting professionals.
Brunnhilde	Siegfried-based characterisation tool for directories and disk images.

bwFLA (emulation as a service)	Browser based, on demand access to emulated software environments.
C3PO	Content profiling tool for preservation analysis
Cfv	Tests and creates checksum verification files.
DD Rescue	Data recovery tool.
DROID	File format identification tool.
EPUB Check	Validation tool for EPUB.
ExifTool	Characterisation and metadata manipulation tool.
FFMpeg	Video characterisation and migration software.
FITS	File format/characterisation.
Flint	A modular and extensible file/format validation framework.
FTK Suite	Software suite for digital forensic professionals.
HAWARP	Hadoop based web archive processing.
Heritrix	Web crawling/archiving.
ImageMagick	Image toolbox.
Isobuster	Data recovery software.
iText	PDF Toolkit for PDF generation, PDF programming, handling & manipulation.
JHOVE2	Format identification/validation/characterisation.
JWAT-Tools	Libraries and tools for reading/writing/validating WARC/ARC/GZIP files
LEAD tools for TIFF	TIFF file format support for .NET (C# & VB, Core, Xamarin, UWP), C/C++, iOS, macOS, Linux, Java, and web developers
matchbox	Near-duplicate image detection.
Md5sum	Tool that calculates and verifies 128-bit MD5 hashes, as described in RFC 1321
MediaConch	Implementation checker, policy checker, reporter, and fixer that targets preservation-level audiovisual files.
MediaInfo	Open-source program that displays technical information about media files, as well as tag information for many audio and video files.
MIXED	Migration to Intermediate XML for Electronic Data) is a service that can convert files with tabular data to supplier independent XML for long term preservation.

HexEditor MX	Hex and disk editor.
NetarchiveSuite	A complete web archiving software package.
Nanite	Scalable identification tool based on DROID.
NZ Metadata Extractor	Preservation metadata extraction/characterisation.
ODF Validator	Open Document Format validator.
Openrefine	A tool for working with messy data: cleaning it; transforming it from one format into another; and extending it with web services and external data.
PDFBox	Command line tools and Java library for working with PDF files.
Pdfapilot (Callus)	pdfaPilot is the solution for conversion of PDF or native documents and emails into robust, searchable PDF or PDF/A files for long-term archiving
PLATO	Web based preservation planning service, Plato is a decision support tool that guides the user through the preservation planning process.
SAXON	Processor for XSLT, XQuery, and XML Schema, including the only XSLT 3.0 conformant toolset.
SIARD I SIARD II	Relational database archiving software.
Siegfried	Signature-based file format identification tool.
ToMaR	Wraps command line tasks for parallel execution as Hadoop map reduce jobs.
TrID	File identification tool.
VirtualDub	Video capture/processing tool.
Webrecorder	A web archiving service to collect and revisit web pages.
Webcurator tool	A tool for managing the selective web harvesting process.
xmllint	Parses one or more XML files, specified on the command line as XML-FILE.

Q16: Community Resources

AV Preserve tools list	A list of tools for the long term preservation of audio visual materials.
COPTR	Community-owned digital preservation tool registry hosted by OPF. COPTR

	collates the digital preservation community's software tool knowledge.
Digipres.org	Community-owned and OPF-hosted digital preservation resources.
Digipres Q&A	Digital preservation and question and answer site hosted by OPF. Anyone can ask digital preservation-related questions to be answered by other members of the community.
DPC Digital Preservation Handbook	A guide to the subject of managing digital resources over time and the issues in sustaining access to them. It will be of interest to all those involved in the creation and management of digital materials.
Inventory of FLOSS (Free/libre open-source software) in the cultural heritage domain	The FLOSS inventory lists all the Free Libre Open Source Software relevant to the cultural heritage world.
JHOVE error message wiki	A list of JHOVE error message compiled for and by the community.
Just Solve the File Format Problem	A Wiki information source on file formats.
Keepers registry	A global monitor on the archiving arrangements for electronic journals.
Library of Congress Recommended Formats Statement	Identifies hierarchies of the physical and technical characteristics of creative formats, both analogue and digital, which will best meet the needs of all concerned, maximizing the chances for survival and continued accessibility of creative content well into the future.
NDSA Levels of Preservation	A tiered set of recommendations for how organizations should begin to build or enhance their digital preservation activities.
OPF Blog	Digital Preservation community blog.
OPF format corpus	An openly-licensed corpus of small example files, covering a wide range of formats and creation tools.
OPF wiki	The OPF wiki gathers knowledge relating to key issues in digital preservation and organises these to allow practitioners, developers and information specialists to better preserve digital heritage.

<p>POWRR Tool Grid</p>	<p>This tool grid is the product of researching digital preservation tools by Digital POWRR team members in early 2013.</p>
<p>Preservieware</p>	<p>An open portal to support the digital preservation community.</p>
<p>PRONOM (format registry)</p>	<p>Online technical registry service hosted by the UK's National Archive.</p>
<p>Wikidata for digital preservation</p>	<p>A technical registry of metadata related to computer software and computing environments captured and collated by the digital preservation community.</p>