

Preservica Enterprise Edition v5.8

Preservica Enterprise Edition is a modular and proven on premise digital preservation platform that can be tailored to meet the needs of different size organizations - from a standard single server configuration to a high through-put multi-server, multi-tenant environment. Additional options include synchronization with catalogue systems, a public access module and a developer's tool kit for work flow customization and tight integration with other systems. Preservica EE is designed for organizations with in-house IT resources, that might also want to keep some content inside their own firewall, as well as be able to combine this with cloud storage (from multiple providers) to create a highly durable and scalable "hybrid" storage environment.

Overview

Preservica EE contains all the functions required to deliver advanced Digital Preservation, as specified in the Open Archives Information System (OAIS) reference architecture (ISO14721:2012): Ingest, Archival Storage, Data Management, Administration, Access, Preservation Planning and Preservation Action.

Content Upload

Ingest allows users to load information into the archive in a managed and efficient manner. Preservica EE provides:

- Locally installable "SIP Creator" to:
 - Build submission packages from locally held files
 - Assign descriptive metadata from fragments created elsewhere or by using a GUI
 - Select where in the hierarchy to place the submission
 - Submit content securely to Preservica
- Upload content directly via browser (no local tool required)
- Create and ingest a website snapshot by crawling a user supplied URL
- Optional package ingest types include:
 - Easily upload content, in multiple formats, resulting from digitization projects
 - Ability to upload DSpace packages and BagIt Bags
 - Ingest ISO disk images (e.g. images of CD, DVD and Hard Disk Drives)
 - Ingest directly from Microsoft SharePoint, PastPerfect and CONTENTdm
 - Preserve email & attachments from Microsoft Outlook (MSG, PST), Gmail and Lotus Notes
- Workflow steps to perform quality assurance including:
 - Virus checking
 - Checksum verification
 - Checking permitted file formats
 - Checking metadata schema compliance
 - Metadata consistency
- Built in technical metadata extraction required for preservation
- Workflow to migrate file formats on ingest using user defined migration pathway (normalization)
- Automated storage processes:
 - Index selected metadata and content
 - Coordinate storage passing the material into the archival database and cloud storage
- Manual and automated workflows to perform all the ingest processes required
- Ability to detect the ingest of duplicate files
- Option to customize Ingest workflows by adding or modifying workflow steps



Archival Storage

- Preservica EE provides the flexibility to store content within a local drive and network drive provided by the customer, and cloud storage in Amazon S3 (fast storage), Amazon Glacier (slow, lower cost storage) and/or Microsoft Azure.
- Two storage adapter types are included with the base license, additional storage adapter types can be added as an option.
 - Simple file system (e.g. Windows or Linux servers, Axiom Pillar)
 - Network Attached Storage / Storage Attached Network
 - Hierarchical storage management(HSM) systems (e.g. FileTek's StorHouse on a Sun server, Oracle SAM-QFS)
- Option to have hybrid on-premise and cloud storage. Different information can be stored in different locations (Amazon S3, Glacier and Microsoft Azure)
- Add support for custom storage types by adding or modifying a storage adaptor
- Options for how information is saved on the drive:
 - Arrange the information physically to mimic the logical hierarchy or keep at a single level
 - Optionally save an additional copy of the metadata to this drive
- Integrity checking of all objects held in storage
 - Checksums calculated on storage for each copy
 - Checksum re-calculated on access or at regular timed intervals on files at rest
 - Confirmation by Preservica EE that when a file is accessed, it is the same as the file loaded in ingest using its own checksum (choose from MD5, SHA-1, SHA-256, SHAR-512)
 - Ability to save selected portions of the collection to an external FTP server

Data Management

Preservica EE contains a range of capabilities to allow users to manage the metadata that describes the content:

- Metadata is held in an internal MySQL database. Alternatively use an existing external Oracle, PostgreSQL, MySQL or SQL Server database
- Descriptive metadata can select from one or more standard metadata schemas provided (display and edit templates included)
 - Encoded Archival Descriptor (EAD) 2002
 - MODS 3.4
 - Dublin Core 1.1
- Alternately, descriptive metadata can make use of user provided descriptive metadata schemas
 - Add XML schema definition
 - Add transforms to present and edit the schema (XSLT)
 - Add templates to allow metadata with default values to be quickly added to information objects
 - Add custom search indexes

- Content management functions provided:
 - Assign one or more metadata fragments to an information object (file, record /deliverable unit, collection)
 - Edit descriptive metadata on any information object
 - Delete information controls with approval steps, editable retention period, alerts and optional time limited recovery
 - View the audit trail of information and digital objects
- Unlimited information hierarchy, with digital objects held at any level
- An optional module allows you to drag and drop to organize and reorganize content and create virtual hierarchies
- Synchronize catalog metadata and collection structure from Axiell CALM, Adlib and ArchivesSpace

Access for Expert Users

Preservica EE provides a Web based "Explorer" to allow the browsing, searching, viewing and downloading of all metadata and content:

- Interactive features:
 - Explore the collection hierarchy, drilling down to information objects of interest
 - Flexible post ingest collection arrangement with drag-and-drop
 - Easy to build Virtual Collections to enable the rapid creation of temporary exhibitions
 - View the files and all metadata (description and audit trail) for any information object
 - View technical metadata for files
 - Thumbnail preview of files (for selected file formats)
 - Download any manifestation (representation) of an entire information object or selected file
 - In-browser rendering to view, listen or play content without having to download the file
- Search tools including:
 - Full text search of permitted file formats and metadata
 - Logical search functions (AND OR NOT etc.)
 - Filtered searches incorporating searching over any registered descriptive schema
 - Limit search to within a selected collection or deliverable unit
- Workflows
 - Launch Access workflows, for example creating a Dissemination Information Package (DIP) for download or delivery, transforming metadata for download into preferred schema either using the provided METS transform or a user provided transform, and sending notifications by e-mail and/or text when the DIP is ready for download
 - Launch Data Management workflows for a specific information object, for example delete or move
 - Customize Access workflows by adding or modifying workflow steps
- Ability to control access to content (or content and metadata) through roles and access control lists (see Administration)



Access for Non-Expert Users

An optional access module can be added to provide a brandable Web based portal to provide advanced search and easy public and internal access to collection content

- Integration with content management systems via CMIS interface (read only)
- Open access for public content or user authentication for more secure content.
- Fast playback of Audio/Visual content

Preservation Planning and Action

- Organization and Planning:
 - Uses File Format Registry based on the UK National Archives PRONOM database
 - Ability to perform preservation planning based on risks stored in the registry, or by simple selection of formats. This includes the ability to select preservation and presentation actions and to select migration pathways
 - Growing list of file formats recognized (more than 1500) and migration pathways (more than 590)
- Edit the Format Registry to add custom formats, migration pathways and tools
- Characterizing the collection:
 - Ability to perform file-level characterization including format identification, format validation and property extraction using best available tool
 - Ability to perform component identification and measurement of component-level essential characteristics (e.g. Web pages)
- Migration
 - Ability to perform migrations including post-migration verification of component-level essential characteristics
 - Ability to create new "Digital Master" preservation copies
 - Ability to create lower quality presentation copies used for dissemination
 - Ability to see and compare pre- and post- migration component information
- Support for Long Term Non-Permanent Records
 - Automatically classify and set access permissions during ingest
 - Automate future actions (e.g. appraise, delete) based on policy rules
 - Search by classification

- Standard tools provided including:
 - Identification / Characterization: DROID 6 .3 and signature file v90, JHOVE 1.16 (JPEG, PDF, HTML, JPEG2000, AIFF, XML, TIFF, GIF, PNG), Java Image IO (for image formats), Apache POI (for MS Office), Entagged Library (for audio formats), FiWalk (for disk images)
 - Migration: for example Native Java Image library (most image formats), FFmpeg (various AV formats)
 - Additional tools: See system requirements
- Ability to perform embedded object extraction and iterative file-level characterization within container formats (e.g. ZIP, GZIP, TAR, MS Office)
- Implementation of all Preservation Actions via the standard workflow engine allowing more complex pathways or actions to be automated and if needed, customized

Administration

- Includes a set of standard reports produced in PDF, XML or CSV formats:
 - Ingest Reports
 - Access Reports
 - Data Management Reports
 - Preservation Reports
 - Administration Reports
 - Proof and authenticity reports
 - Storage & file type reports
- Create and Upload custom reports
- Administration
 - Maintain Transfer Agreements with content suppliers
 - Manage schemas used for content description
 - Fast file existence check
- User Management & Authentication
 - Create and suspend users
 - Create roles and assign to users
 - Users requested new password
- Comprehensive security model for information objects
 - Define tags saying which actions can be performed by which user roles
 - Assign tags to collections and records (deliverable units) allowing fine grain user access control to information objects



Support

Operations

- Preservica EE is installed and configured on client hardware, on either Linux or Windows operating systems, by Preservica staff
- Configured to use a client's Open LDAP, Active Directory, or SAML provider (e.g. Azure AD) for access control
- Configured to use a client's database. MySQL, Microsoft SQL Server, PostgreSQL, or Oracle are supported
- Training is included for users and managers of the system

Professional Support

Preservica provides a full ongoing support service, advice and guidance, and bug fixing.

Development Program

- Preservica EE is under active development with major and minor releases made throughout the year. All development is performed under our ISO 9001 Quality Management System
- EE has an optional Software Development Kit (SDK) which allows development partners to customize and integrate with other systems.

System Requirements

- Browser Access
 - Chrome
 - Internet Explorer v11
 - Firefox
- Client platforms
 - SIP Creator: Windows or Linux desktops
- Storage
 - See section for Archival Store
- Database Platforms
 - MySQL 5.6 & 5.7
 - Microsoft SQL Server 2008 R2 & 2014
 - PostgreSQL 9.5, 9.6
 - Oracle 11gR2, 12c
- Preservica EE Operating Systems (all 64bit)
 - Windows Server 2008r2, 2012
 - SUSE Linux Enterprise Server v12
 - Red hat Linux v6, v7
 - CentOS Linux v6, v7
- Characterization & migration tools (open source)
 - MedialInfo
 - Libre Office
 - iCat
 - ImageMagick
 - DCRaw
 - Ghostscript
 - wkhtmltopdf
 - HandbrakeCLI