

# Preservica Cloud Editions v6.5

Preservica is a world leader in digital preservation technology and research. Our active preservation solutions are used by leading businesses, archives, libraries, museums and government organizations globally, to safeguard and share valuable digital content, collections and electronic records, for decades to come. Preservica Cloud Edition Essentials (CEE) and Cloud Edition Professional (CEP) are comprehensive cloud-hosted Software as a Service active preservation and public access platforms. CEE and CEP are built to OAIS standards and combine a rich set of preservation workflows with affordable and durable cloud storage. CEE and CEP are most suited for small and mid-sized organizations with limited local IT resources or servers, looking for a comprehensive out-of-the-box preservation solution, optimized for smaller or medium volume collections. CEE and CEP are designed to support operations in regulated industries and the GDPR.

## Overview

Preservica CEE and CEP contains all the functions required to deliver advanced Digital Preservation, as specified in the Open Archives Information System (OAIS) reference architecture (ISO 14721:2012): 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 CEE and CEP features include:

- Upload and prepare content directly via web browser (no tool download required):
  - Simple drag and drop interface for individual files up to browser limit (approx. 4GB)
  - Upload larger files via a secure holding area
  - Immediately safeguard content to 99.999999999% durability while you prepare your content for upload into Preservica
  - Easily re-arrange content hierarchy before upload
  - Add, edit and enrich descriptive metadata
  - Restore autosaved temporary projects at any time
- An optional Bulk Upload Service enables uploads of up to 80TB at a time quickly and easily, working in conjunction with AWS Import/Export Snowball
- Create and ingest a website snapshot by crawling a user supplied URL
- Package ingest types include:
  - Easily upload content in multiple formats, born digital or from digitization
  - 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 PastPerfect and CONTENTdm
  - Preserve email and MBOX files from Microsoft Outlook (MSG, PST), Gmail and Lotus Notes
  - Compressed & uncompressed folders containing common formats for documents, AV files, and more
- Workflow steps to perform quality assurance including:
  - Virus checking
  - Checksum verification
  - Checking permitted file formats
  - Checking metadata is well-formed
- Built in technical metadata extraction required for preservation (for more see Preservation below)
- 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 required ingest processes
- Ability to detect the ingest of duplicate files
- Configure and automate the capture and ingest of social media accounts using Preservica's Intelligent Archiving Engine



## Archival Storage

Preservica CEE or CEP stores its content within Amazon Web Services (AWS) storage systems. This provides the following functions:

- Use of AWS Simple Storage Service (S3) to hold digital assets (files) for fast, frequent access
- Save files to multiple servers in multiple data centers within a given region
- Ability to split collections between Amazon S3 (fast storage) and Amazon S3 Glacier (slow, lower cost storage)
- Encrypt content stored in AWS
- Optional isolation of content in S3 and Glacier to store the content for one organization only
- 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 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 secure FTP server
- Optionally maintain additional backup copies of files and metadata in a second supported cloud storage location [customer account] - CEP only.

## Data and Collection Management

Preservica CEE and CEP contain a range of capabilities to allow users to manage the metadata that describes the content:

- Metadata is held in Amazon Web Services' Relational Database Service
- 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 assets
  - Add custom search indexes
- Content management functions provided:
  - Assign one or more metadata fragments to an asset or folder
  - Edit descriptive metadata on any asset post-ingest
  - Make bulk changes to the metadata of multiple assets. Includes replacing, clearing or adding a fragment
  - View the audit trail of information and digital objects
- Unlimited information hierarchy, with digital objects held at any level
- Drag and drop to organize and reorganize content and create virtual hierarchies
- Synchronize catalog metadata and collection structure from Axiell CALM, Adlib and ArchivesSpace

- Classify personal data information for the GDPR fragment
- Create reference metadata to refer to the same information across multiple assets
- Automated retention and disposition of records – CEP only:
  - Meet regulatory, statutory and compliance requirements for long-term records
  - Create and apply retention policies to single or groups of records
  - Automate disposition with records presented for appraisal or deletion
- Replace incorrect content within an asset without having to re-upload

## Access for Expert Users

Preservica CEE and CEP provide 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 create linked content to enable the rapid creation of temporary exhibitions
  - View the files and all metadata (description and audit trail) for any asset
  - View technical metadata for files
  - Thumbnail preview of files (for selected file formats)
  - Download any asset or selected part
  - In-browser rendering to view, listen or play content without having to download the file including images, office formats, emails, social media posts, web pages, captioned videos, and the rendering of 3D assets
  - Download assets and folders in OPEX (Open Preservation Exchange) format for enhanced usability
- Search tools including:
  - Full text search of permitted file formats and metadata
  - Logical search functions (AND OR NOT etc.)
  - Fielded searches incorporating searching over any registered descriptive schema
  - Limit search to within a selected folder
  - Export search results into CSV for custom reporting on metadata
- 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 asset, for example delete or move)
  - Optional workflow to migrate file formats on access using user defined migration pathways (normalization)
- Ability to control access to content (or content and metadata) through roles and access control lists (see Administration)



## Access for Non-Expert Users

- An easy to customize web based portal to provide advanced search and controlled public and internal access to content
- Range of intuitive, customizable themes
- Open access for public information or user authentication for more secure content
- Fast playback of Audio/Visual content
- Integration with content management systems via CMIS interface (read only)
- Accessibility WCAG 2.0 AA (excluding optional animated theme)
- Display most popular assets and allow users to compile favorites
- Enable restricted or unrestricted downloading of assets

## 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 simple selection of formats, including the ability to select preservation and presentation actions and to select migration pathways
  - Growing list of file formats recognized (more than 1700) and migration pathways
- Characterizing the collection:
  - Ability to perform file-level characterization including format identification, format validation and property extraction
  - Ability to perform component identification and measurement of component-level essential characteristics
- Migration:
  - Ability to perform migrations including postmigration 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
  - Migrate files to Microsoft Office Open XML documents (.docx), in line with the NARA Policy recommendation
  - Ability to migrate multi-part assets such as multi-page TIFF to multiple JPEGs and PDF to multiple JPEGs. This allows all pages to be rendered in the multi-image gallery renderer and enhances the viewing experience
- Standard tools provided including:
  - Identification / Characterization: DROID 6.3 and signature file v94, JHOVE 1.20 (JPEG, PDF, HTML, JPEG2000, AIFF, XML, TIFF, GIF, PNG), Native XML (validation only), Media Info (for audio and video formats), FiWalk (for disk images), ExifTool (for MS Office OLE formats and other image formats e.g. Canon RAW), Apache POI (for MS Office OpenXML formats validation only), Native Java (for OpenDocument formats and MS Office OpenXML formats property extraction only)
  - Migration: Libre Office (various office formats to OpenDocument formats, PDF, PDF/A and various vector graphics formats to SVG and PDF), Imagemagick (most image formats), FFMPEG (various AV formats), Handbrake (for video)

- Implementation of all Preservation Actions via the standard workflow engine allowing more complex pathways or actions to be automated
- Ability to migrate individual image and video files in a multi-part asset, such as a social media record, to new formats over time to ensure the record can always be read in full

## 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
  - Information with GDPR personal data classification 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
  - Option to enable two factor authentication, compatible with common authentication applications
- Comprehensive security model for assets
  - Define tags saying which actions can be performed by which user roles
  - Assign tags to assets and folders allowing fine grained user access control to information objects
- Optional integrated in-product assistance and step-by-step guidance



## Support

### Operations

- Preservica CEE and CEP are operated by a professional operations team that keeps the system operating, maintains backups, checks performance and does test restores
- The hardware platforms are provided by Amazon Web Services via locations in US (N. Virginia), Canada (Montreal), EU (Dublin) or Australia (Sydney)
- Operations fall under our ISO 27001 (Information Security Management System) accredited processes
- There is no system administration requirement for the customer
- Training is included for users and managers of the system
- Remove your data from Preservica with no cost exit with no data lock-in

### Professional Support

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

### Integration

Preservica CEE and CEP include extensions for use and integration:

- Synchronization with supported catalogs (See Data Management section)
- OAI-PMH API for providing metadata harvesting information
- Optional REST API for content read and write actions up to documented transaction limit - CEP only.

### Development Program

Preservica CEE and CEP are under active development with major and minor releases made throughout the year. All development is performed under our ISO 9001 Quality Management System.

### System Requirements

Access to Preservica is provided through the following supported browsers:

- Google Chrome v77+
- Internet Explorer v11
- Mozilla Firefox v69+
- Safari v12 (web browser upload only)

An open connection to the internet is required for connection to Preservica CEE and CEP.