



Ex Libris Rosetta

A Complete Digital Asset Management and Preservation System



Product Description

CONFIDENTIAL INFORMATION

The information herein is the property of Ex Libris Ltd. or its affiliates and any misuse or abuse will result in economic loss. DO NOT COPY UNLESS YOU HAVE BEEN GIVEN SPECIFIC WRITTEN AUTHORIZATION FROM EX LIBRIS LTD.

This document is provided for limited and restricted purposes in accordance with a binding contract with Ex Libris Ltd. or an affiliate. The information herein includes trade secrets and is confidential.

DISCLAIMER

The information in this document will be subject to periodic change and updating. Please confirm that you have the most current documentation. There are no warranties of any kind, express or implied, provided in this documentation, other than those expressly agreed upon in the applicable Ex Libris contract.

Any references in this document to non-Ex Libris Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this Ex Libris product and Ex Libris has no liability for materials on those Web sites.

Copyright Ex Libris Group, 2017. All rights reserved.

Documentation produced January 2017

Web address: <http://www.exlibrisgroup.com>

Index

Index	3
Executive Summary	4
Product Information	8
Adherence to Metadata Standards.....	9
System High-Level Architecture	11
Ingest.	12
Working Area	13
Permanent Repository	15
Operational Repository.....	15
Management	16
Administration	16
Preservation	17
Delivery and Access	19
Scalability.....	21
Openness.....	22
Business Logic-Driven Workflows	23
Rosetta Community-Driven Knowledge Base	24
Find Out More.....	24

Executive Summary

International research has shown that the “digital universe” is growing at an unforeseen rate. Indeed, current forecasts expect the volume of digital content—from CCTV footage and personal photographs to academic research, research output in digital formats, Web sites, and corporate financial records—to double in size every 18 months¹.

As the options for creating and duplicating digital items multiply, comparable options for storing those items remain limited. Documents, images, and film produced in recent decades will be inaccessible to future generations as a result of continuing waves of innovation that render older storage systems obsolete.

As described in a white paper from Preservation and Long-term Access through Networked Services (Planets), “The more information we produce, and the more we can hold in a given space, the shorter the time we seem to be able to keep hold of it for.”²

Libraries, museums, and archives face particularly stark realities in this digital environment. Not only do they continue to store and provide access to the thousands of objects and resources in their care, but they must ensure that the huge quantities of digital items being acquired with every passing month are stored, preserved for the long term, and available for future generations.

With increasing numbers of resources that are born digital, rather than being digitized from a physical copy, the challenge for “memory” institutions (libraries, museums, and archives) is unique. Whereas the records of governments, companies, or agencies might be expected to become outdated and irrelevant over time, memory institutions are committed to preserving the records in their care in perpetuity.

¹ John F. Gantz & David Reinsel, “As the Economy Contracts, the Digital Universe Expands,” an IDC Multimedia White Paper sponsored by EMC, May 2009.

² Pauline Sinclair, “The Digital Divide: Assessing Organisations’ Preparations for Digital Preservation,” a Planets (Preservation and Long-term Access through Networked Services) white paper, March 2010.

A study by Planets (Preservation and Long-term Access through Networked Services)—a European Union-funded project to assess digital preservation needs—found that while 80% of organizations currently store documents and images, by 2019 over 70% will need to preserve databases, web sites, audio files, and video files, as well³. The emerging task for libraries is not only how to store the huge diversity of digital assets being generated but also how to equip institutions to scale their preservation efforts to meet the rapidly expanding volume of data.

Even experts have not accurately predicted the speed at which the digital universe is expanding. Research by the International Data Corporation (IDC) in 2010 underestimated the previous year's rate of growth by 3%, while their earlier paper in 2008 misjudged the global volume of digital items by 10%⁴. As institutions look to a future of absorbing millions of digital objects each year, scalability is a fundamental requirement for their digital preservation architecture.

In the last three decades digital technology has enabled us to create, use, and be enriched by information in ways that were unthinkable a generation ago. The growth in the number of digital items in today's library collections—items that have undergone digitization and items that were "born" digital—has led to an understanding that new actions must be taken to preserve these digital assets and make them available to future generations.

The challenge of preserving digital material is most acute regarding items that were created in digital format—that is, "born" digitally. The vast majority of this material exists exclusively in digital format, a fact that makes the preservation of digital information critical to the perpetuation of our cumulative knowledge.

While many organizations have systems in place for storing and managing digital objects, these systems are not always designed with preservation in mind and include the proper functions and features to facilitate digital preservation. Digital preservation is about guaranteeing the continued usability of and access to digital content tomorrow and well into the future. Digital asset management systems and digital repositories focus on facilitating the day-to-day use of digital content, whereas a digital preservation system offers discovery and access options,

³ Pauline Sinclair, "The Digital Divide: Assessing Organisations' Preparations for Digital Preservation," a Planets (Preservation and Long-term Access through Networked Services) white paper, March 2010.

⁴ John F. Gantz et al., "The Diverse and Exploding Digital Universe: An Updated Forecast of Worldwide Information Growth through 2011," an IDC white paper sponsored by EMC, March 2008.

functionality and workflows for ingesting materials, ongoing risk analysis, and the continual integrity of stored items. Although preservation focuses on risk management, we would be mistaken if we equated preservation with backup or disaster recovery.

In today's economic environment it is clear that any investment made in digitization and storing of digital content should be accompanied with a strategy for sustainable digital preservation, one that will allow institutions to not only ensure short-term access but also provide long-term solutions for the sustainability of the digital content.

Quoting Jay Schafer, Director of Libraries at UMass Amherst, "The role of libraries in the print world—collecting, providing access to, and preserving our cultural heritage—does not change as we move into a digital realm. The role of libraries remains the same but we must take on new challenges in determining how to fulfill this role"⁵. It is clear that preservation as a service the library provided and still provides for its physical collection must make the transition into the digital world as more and more national, state and academic libraries are transitioning from traditional paper-based libraries into digital libraries.

As the digital preservation matures, an emerging standard appear to set the standard for digital preservation and archiving solutions. The Open Archival Information System (OAIS) reference model describes characteristics of a digital preservation system⁶. The model has become widely accepted among preservation bodies and experts worldwide and has been used as a guideline to evaluate current implementations of preservation and archiving initiatives⁷. The OAIS model specifies six high-level functions that must be present in a digital preservation system:

- Ingest
- Storage
- Data management
- Administration
- Preservation planning
- Access

⁵ <http://www.library.umass.edu/assets/aboutus/Reports/SPR-SUMM-2011-Newsletter-41CMS.pdf>

⁶ <http://public.ccsds.org/publications/archive/650x0b1.pdf>

⁷ See, for example, *Assessment of UKDA and TNA compliance with OAIS and METS Standards*, at http://www.jisc.ac.uk/uploaded_documents/oaismets.pdf.

These six functions are indeed an integral part of Ex Libris Rosetta for digital libraries, whose release was announced in January 2009. Developed in partnership with the National Library of New Zealand and reviewed by a peer review group of world-renowned preservation experts, Ex Libris Rosetta addresses libraries' need to collect, manage, archive and preserve a wide variety of digital objects in different formats and structures.

Ex Libris Rosetta enables libraries to manage digital entities end to end - from submission to dissemination. A rule-based workflow engine and open architecture allow institutions using the system to develop unique plug-in tools and applications to enhance the system's ingest, management, preservation, and delivery processes.

Already in use by dozens of institutions worldwide, Rosetta includes a unique content preservation planning and action module which provides institutions the means to identify format risks, evaluate mitigation alternatives and execute preservation actions to ensure digital content's longevity and access. The module relies on a global knowledge base, storing information on formats, applications and more and helps institutions to share information about preservation related issues.

Product Information

Released for general availability in January of 2009, Ex Libris Rosetta included a rich set of features allowing libraries, archives, museums and other memory institutions to collect, manage, archive and deliver digital information of different types. With an emphasis on workflow optimization and easy-to-use web based user interface Rosetta is already used to preserve digitized and “born” digital cultural heritage, research data, publications and more.

As a development partner, The National Library of New Zealand was the first Rosetta site to go live with version 1.0 of the product in October 2010. Since then Rosetta version 2.0 was released in May 2010 and included the unique preservation planning and action module as well as the system knowledge base as well as number of other significant new developments. Version 3 has introduced an improved UI, Collection Management, a new search mechanism and much more. Rosetta version 4, which was released in mid-2014, included major improvements in system scalability, performance and robustness, new DAM functionalities and much more. Rosetta version 5 was released in December 2015 and included two-way metadata synchronization with external Content Management Systems, enhanced support for Research Data, integration with Google Scholar and much more. In 2016 Rosetta received a new UI, new viewers and major additions throughout the system. Rosetta’s 2017 roadmap includes exciting features in both DAM, preservation and research data, such as SAML support, BagIt submission, native IIIF support and much more.

Adherence to Metadata Standards

One of the key guiding principles during the development of Ex Libris Rosetta was to ensure the solution adheres to international standard for digital preservation. In addition, it was crucial to ensure that the system data model will sufficiently flexible to manage digital content of different types, including: images, text document, audiovisual and more. The goal was to create a data model that will:

- Support a structure and metadata for long-term digital preservation
- Support a variety of formats and materials in one data model
- Support a variety of institution types (museums, libraries, and archives) with preservation needs
- Conform to standard and open practices

Rosetta's data model is based on the Preservation Metadata: Implementation Strategies (PREMIS) working group's metadata elements and the Metadata Encoding and Transmission Standard (METS):

- PREMIS as the conceptual model outlining the entities and metadata required for preservation
- METS as the wrapper for encoding descriptive, administrative, and structural metadata

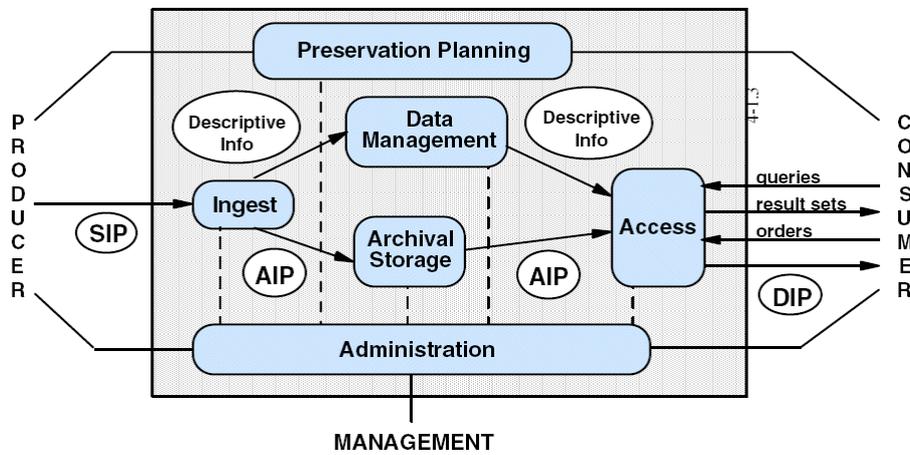
The METS standard was already commonly used for digital repositories and provided flexibility as a container in structure and storing metadata sufficient for digital preservation. The PREMIS metadata elements are stored in the Rosetta METS container in three different levels; each can hold a single or multiple entries for the level below:

- **Intellectual Entity:** a coherent set of content that is reasonably described as a unit, for example, a particular book, map, photograph, or database
- **Representation:** set of files, including structural metadata, needed for a complete and reasonable rendition of an Intellectual Entity
- **File:** named and ordered sequence of bytes that is known by an operating system

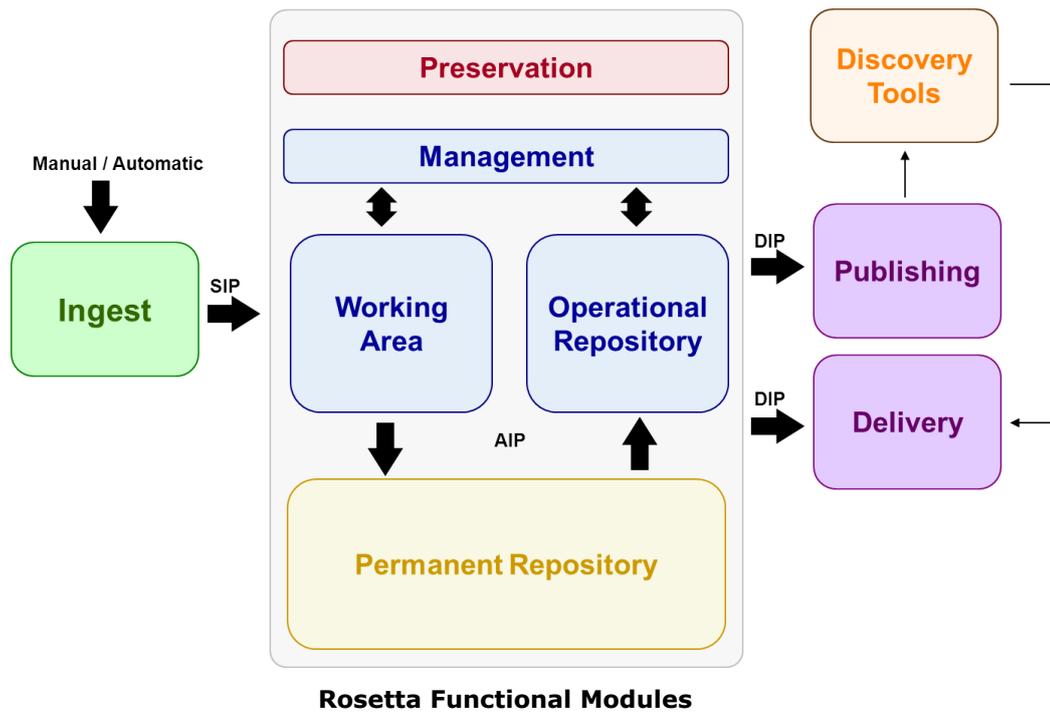
A preservation system must also document and preserve the descriptive information of its digital entities, as well as provide the ability to integrate with existing and future cataloguing solutions. To support this, Rosetta uses the extendable Dublin Core elements to store descriptive metadata in the system as the main metadata standard. Additional metadata standards are supported in the system as source metadata that can be included as part of the submitted package (SIP), stored and edited for each intellectual entity as well as published and delivered to end users.

System High-Level Architecture

According to the OAIS reference model, the functions that are mandatory for any preservation system are ingest, archival storage, data management, administration, preservation planning, and access⁸.



In conformance with the OAIS model, Rosetta consists of the following modules: Ingest, Working Area, Permanent Repository, Operational Repository, Management, Delivery, Publishing, and Preservation.



⁸ Functional entities of OAIS reference model, reprinted from *Reference Model for an Open Archival Information System (OAIS)*, Consultative Committee for Space Data Systems, January 2002

Ingest

The ingest module enables external producers (publishers, end users) or internal producers (institution staff) to upload and save submission information packages (SIPs)⁹ for ingestion into the preservation repository. The module supports multiple workflows—automatic, semiautomatic, and manual—which can be assigned to the producer depending on the producer’s agreement with the institution. The ingest module provides producers with a web-based user interface through which they can upload files, provide metadata, and define access restrictions. Files and metadata can also be uploaded to an FTP site from which the ingest module downloads them. Additionally, a software development kit (SDK) provides seamless integration of the ingest module with record-management systems, thus facilitating the process of making large, frequent deposits.

A SIP can contain one or more intellectual entities¹⁰. An intellectual entity can be either simple (a single digital entity or file to be managed and preserved) or complex (multiple digital entities or files that are managed and preserved as one or more groups). The ingested content is wrapped as a submission information package (SIP) and sent to the working area.

⁹As explained in the draft *Submission Information Package (SIP) Specification* version 1.0, “under the OAIS model, material from a content provider is transmitted to the archive in a form called a Submission Information Package (SIP). The archive Ingest function accepts the SIP and potentially transforms its contents into an internal form called an Archival Information Package (AIP) for long-term preservation”.

¹⁰As defined in the PREMIS Data Dictionary for Preservation Metadata, an intellectual entity is “a set of content that is considered a single intellectual unit for purposes of management and description: for example, a particular book, map, photograph, or database. An Intellectual Entity can include other Intellectual Entities; for example, a Web site can include a Web page; a Web page can include an image. An Intellectual Entity may have one or more digital representations” (<http://www.loc.gov/premis/v2/premis-2-0.pdf>).

ExLibris Rosetta Deposit Ex Libris Demo Producer John Smith Demo Institution English

Home Producers

Deposit Activities

Draft (2) Submitted (272) Returned (1) Declined (4) Approved (26)

Add Deposit Activity Find: [] In: All Go

21 - 40 of 272 Records

Deposit ID	SIP ID	Title	Material Flow Type	Created On	Submitted On	View	Notes	History
21	13532	1502	Digitized Manuscript	Unpublished	06/02/2017 14:57:16	06/02/2017 15:01:40	View	Notes History
22	13499	1500	Van Gogh Images	Serial	30/01/2017 23:52:00	30/01/2017 23:52:00	View	Notes History
23	13498	1499	Marketing movie	Serial	30/01/2017 23:51:41	30/01/2017 23:51:40	View	Notes History
24	13497	1498	Harvested website	Serial	30/01/2017 23:51:20	30/01/2017 23:51:20	View	Notes History
25	13496	1497	Database	Serial	30/01/2017 23:51:00	30/01/2017 23:51:00	View	Notes History
26	13495	1496	Baght bag	Serial	30/01/2017 23:50:43	30/01/2017 23:50:42	View	Notes History
27	13494	1495	Sound Archive	Serial	30/01/2017 23:50:20	30/01/2017 23:50:20	View	Notes History
28	13493	1494	Research Data	Serial	30/01/2017 23:50:05	30/01/2017 23:50:05	View	Notes History
29	13492	1493	Photographs	Serial	30/01/2017 23:49:44	30/01/2017 23:49:43	View	Notes History
30	13472	1492	University Collections	Unpublished	27/01/2017 16:33:42	27/01/2017 16:39:20	View	Notes History
31	13459	1491	University Archives	Serial	23/01/2017 23:52:06	23/01/2017 23:52:06	View	Notes History
32	13458	1490	Books	Serial	23/01/2017 23:51:45	23/01/2017 23:51:45	View	Notes History
33	13457	1489	Journals	Serial	23/01/2017 23:51:25	23/01/2017 23:51:24	View	Notes History
34	13456	1488	Newspapers Collection	Serial	23/01/2017 23:51:05	23/01/2017 23:51:05	View	Notes History
35	13455	1487	Musical Notes	Serial	23/01/2017 23:50:45	23/01/2017 23:50:45	View	Notes History
36	13454	1486	Historical Photos	Serial	23/01/2017 23:50:25	23/01/2017 23:50:25	View	Notes History
37	13453	1485	Mixed collections	Serial	23/01/2017 23:50:04	23/01/2017 23:50:04	View	Notes History
38	13452	1484	Mozart Historical Recordings	Serial	23/01/2017 23:49:44	23/01/2017 23:49:43	View	Notes History
39	13432	1483	Biological Research	Unpublished	18/01/2017 11:18:28	18/01/2017 11:21:52	View	Notes History
40	13419	1482	State Archives	Serial	16/01/2017 23:52:04	16/01/2017 23:52:04	View	Notes History

21 - 40 of 272 Records

ExLibris Terms of Use © Ex Libris, a ProQuest Company 2017

Ingest Module: Web Deposit User Interface

ExLibris Rosetta Deposit Ex Libris Demo Producer John Smith Demo Institution English

Home Upload Files

Please upload your files here:

Select File

Files for upload

Add more files, add folder or remove all files

Change view: Thumbnails

Total files: 13

Upload

Selected Files

Back Cancel Next

Ingest Module: File Upload Interface

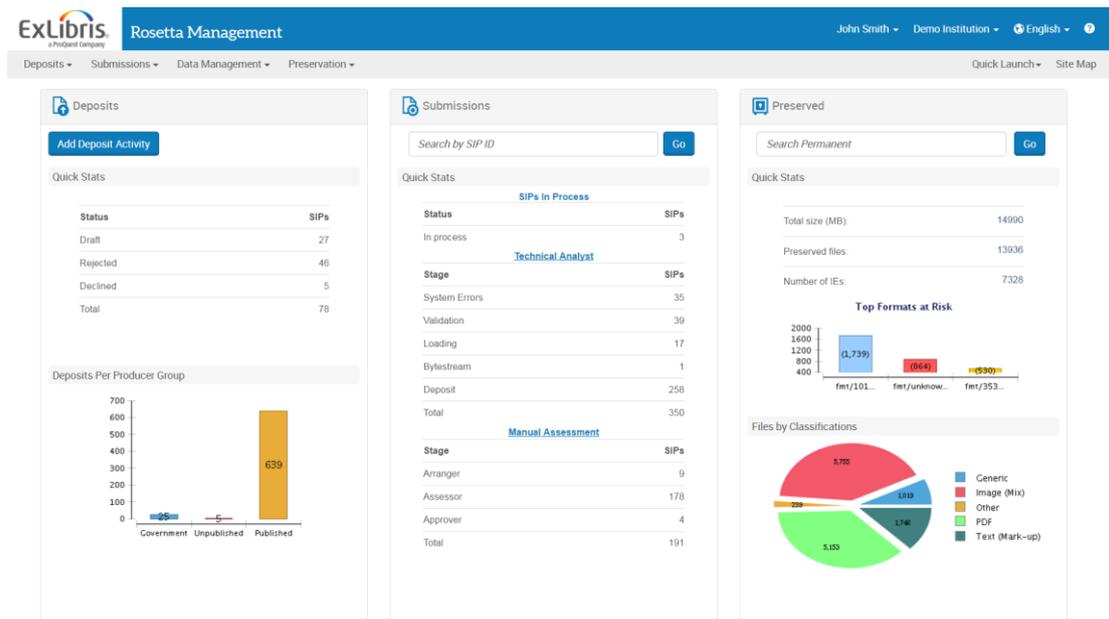
Working Area

The working area is a repository for managing internal ingests processes for materials that have been submitted to the system. It represents the point at which a SIP becomes an archival information package (AIP) in the OAIS model.

The working area supports both automatic and manual processes, such as the following:

- Virus checks, checksums, format identification, technical metadata extraction and validation to ensure the integrity of the SIP data stream (files) and metadata
- Curatorial and accession processes that monitor the quality of the submitted material to verify that it is complete and not damaged
- Enrichment processes that convert information into a stable format when necessary and enrich objects with relevant metadata

Some of the actual processes (for example, virus checks and format validation) are carried out by means of third-party tools, such as designated user-interface tools and APIs. The working area's workflow engine facilitates the embedding of such tools.



Working Area Module: Dashboard

Permanent Repository

The permanent repository module is the long-term preservation repository for approved and successfully enriched materials. Its core function is to store information in perpetuity. Items are stored under a write-once, read-only policy with full disk replication to ensure that data and objects are completely secure.

Full disk replication provides the following benefits:

- The system's dependence on the database is minimal, because everything can be reconstructed from the repository disk.
- Information about each object is self-contained and includes full, relevant metadata for managing preservation actions over time.
- The system's dependence on software applications is kept to a minimum because the disk structure is open and is documented according to industry standards
- Objects stored in the permanent repository that have undergone changes through preservation actions are recorded as a version with auditing information
- Migration and media refreshment are easy to implement.

To ensure the long-term integrity of the information stored in the permanent repository, virus and fixity checks are run directly on the items stored in it.

Operational Repository

Besides the permanent repository, Rosetta provides an operational repository that enables searching, indexing, and quick access. The operational repository is where preservation actions take place. AIP that are stored in the permanent repository and are identified as candidates for preservation action by the preservation planning module are subjected to the requisite preservation action in the operational repository. The output of the preservation activity is re-ingested into the permanent repository and stored as a new version of the original intellectual entity. An example is a preservation action, initiated by staff members, for migrating image files from TIFF to JPEG 2000 format. First, the system searches the permanent repository and copies to the operational repository the AIP that includes the TIFF images. After the new images are created, they are moved to the permanent repository as a new version of the same AIP.

The screenshot displays the Rosetta Management interface. At the top, the ExLibris logo is on the left, and the user 'John Smith' is logged in. The main navigation bar includes 'Deposits', 'Submissions', 'Data Management', and 'Preservation'. The current page is 'Data Management: Search for Objects / Details' for Intellectual Entity ID IE44463. A table shows metadata: Created on 23/10/2013 21:23:09, Created by admin1, Committed on 16/11/2016 00:56:09, Committed by admin1, SIP ID 679, and Version 10. Below this, a list of files is shown under 'IE (IE44463)', including 'Preservation Master Revision 1 (REP44464)' with three files and 'Modified Master Revision 1 (REP44468)' with three files. A 'View Object' button is present. To the right, a 'View Object' summary table lists: PID IE44463, SIP ID 679, Created on 23/10/2013 21:23:09, Created by admin1, Updated on 19/06/2016 15:25:49, Updated by admin1, Entity Type ALTO, Number of Reps 3, and Status ACTIVE.

Operational Module: Intellectual Entity Editor

Management

Rosetta has a Web-based interface for managing the entire preservation system as well as the digital records. The Management module provides functionality for maintenance tasks, metadata editing, searching, and other tasks.

Administration

To support system management, the administration module includes the following cross-application functionalities, and more:

- **System configuration:** Users manage system configuration through a back-office Web-based, wizard-like interface.
- **Reporting:** Integrated into Ex Libris Rosetta is the open source, Eclipse-based Business Intelligence and Reporting Tools (BIRT) reporting system, which provides detailed lists and statistics on various aspects of the system and collections. Alternatively, the institution can choose to integrate an external reporting system, such as Crystal Reports.
- **User management:** Staff users, patrons, and producers can each have a set of customizable roles and privileges. The system supports integration with external identification, authentication, and management (IAM) systems via standard interfaces such as LDAP.
- **Auditing:** All activities performed on the digital objects maintained in the system are fully logged and traceable. Therefore, any changes that are made

to an object after its ingestion into the system are recorded. Action information is stored in the system according to the effect that the action might have on the stored object.

- **Monitoring:** The system includes an interface that enables staff to monitor maintenance jobs and other processes that are running in the system.

Active	Plug-In Type Name	Plug-In Name	Description	Contact Person	Plug-In Version			
1	MDEXtractorPlugin	AIFF-hul	AIFFHULMDEXtractorPlugin	Exlibris Ltd	2.2	View	Edit	Delete
2	MDEXtractorPlugin	AIFF-hul-1.10	AIFFHULMDEXtractorPlugin	Exlibris Ltd	3.0	View	Edit	Delete
3	MDEXtractorPlugin	ASCII-hul	ASCIIHULMDEXtractorPlugin	Exlibris Ltd	2.2	View	Edit	Delete
4	MDEXtractorPlugin	ASCII-hul-1.10	ASCIIHULMDEXtractorPlugin	Exlibris Ltd	3.0	View	Edit	Delete
5	MDEXtractorPlugin	BYTESTREAM-hul	BYTESTREAMHULMDEXtractor	Exlibris Ltd	2.2	View	Edit	Delete
6	MDEXtractorPlugin	BYTESTREAM-hul-1.10	BYTESTREAMHULMDEXtractor	Exlibris Ltd	3.0	View	Edit	Delete
7	FieldValidatorPlugin	ChecksumDigitValidatorPlugin	RegexpFieldValidator	Exlibris Ltd	3.0	View	Edit	Delete
8	ConverterRegistryPlugin	CMS-converter	CmsConverterPlugin	Exlibris Ltd	6.0	View	Edit	Delete
9	PIGenerator	CMSGenerator	CMS sequence Generator	Exlibris Ltd	4.0	View	Edit	Delete
10	RepositoryTaskPlugin	DCReplacePlugin	Finds and replaces a DublinCore field value	Exlibris Ltd	2.0	View	Edit	Delete
11	DecomposerPlugin	decomposeArcFile	decomposeArcFile	National Library New Zealand	2.0	View	Edit	Delete
12	RepositoryTaskPlugin	DummySleepTaskPlugin	Sleeping for a desired time	Exlibris Ltd	1.0	View	Edit	Delete

Administration Module: Plug-in Configuration Menu

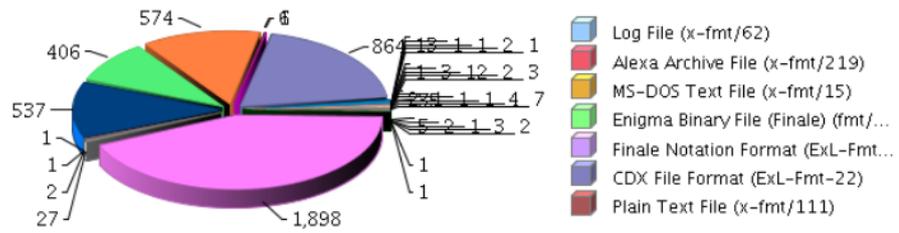
Preservation

The OAIS reference module specifies that a preservation system should include a preservation planning module. Rosetta includes a preservation planning module that comprises a format library¹¹, risk analysis, evaluation, and preservation actions. The module allows institutions to manage and carry out the full preservation process of identifying a risk, selecting the best among possible alternative solutions, and testing/activating the preferred preservation process.

- **Risk analysis:** Based on format risks the system assesses the possibility of a format's becoming obsolete and provides a report on the risk to the repository manager. Risks can be added by the institution, but institutions can view risk analyses that derive from other institutions as well.

¹¹ A list of formats with relevant metadata on the different formats. The format library is based on PRONOM, a global format library maintained by the United Kingdom National Archives.

Formats at Risk



Total number of files at risk: 4422

#	Risk	Format Name	Format ID	Number of Files	Percentage
1	Zero Applications	Extensible Markup Language	fmt/101	1898	13.04%
2	Zero Applications	Tagged Image File Format	fmt/353	864	5.94%
3	Zero Applications	JPEG 2000	X-fmt/392	574	3.95%
5	Bad Compression Scheme	JPEG 2000	X-fmt/392	573	3.94%
7	JP2 - Unknown ColourSpace	JPEG 2000	X-fmt/392	567	3.90%

Risk report based on the files in the repository and the registered risks

- Test bed:** The institution can compare a number of preservation path options for materials that are identified as high risk. The system tests each of these options on a sample set of items and allows the institution to rate the success of the option based on parameters the institution has defined. Institutions can also view preservation plans from other institutions.

Rosetta Management

John Smith | Demo Institution | English

Deposits | Submissions | Data Management | Preservation
Quick Launch | Site Map

Preservation: Preservation Plans / Details

Plan Name	fmt/7 Preservation Plan (TIFF)	Creation Date	22/03/2012	Created By	admin1
Source Format	fmt/7	Source Classification	Image (Mix)	Risk Code	TIFF Low Compression Scheme
Preservation Set	8131949	Test Set	8131950	Sustainability Factors	Comparison Report

Alternative Evaluation Results

	BMP Migration	TIFF to JP2000	Uncompress TIFF files
Cost per Day	Cheap	Expensive	Reasonable
Cost-Software > Initial	50000	50000	10000
Integrity > File Format Verification	yes	yes	yes
Integrity > Traceability of change	no	no	yes
opinion	This would be very expensive from a storage point of view	Valid option. Not ideal though	V good option, keeping the file format

Back
View as PDF

Review of the automatic comparison of the migration alternative

- **Preservation action:** After the institution has determined the best preservation approach for a given set of items, the system implements the preservation action on the entire set. Each preservation action that is carried out is labeled with an ID and a description. Information regarding objects that are undergoing or have already undergone the preservation activity is stored in the preservation planning module. Thus, staff users can view the status of a preservation action and automatically run the same action on additional objects of the same format.

Delivery and Access

Dissemination information packages (DIP) are created in the access module. These packages are used for the dissemination of information stored in the permanent repository.

The access module consists of two components:

- The publishing component enables external systems (for example, search engines and resource discovery solutions such as Primo[®] from Ex Libris) to use standard communication protocols, such as OAI-PMH, Search/Retrieve via URL (SRU), and Search/Retrieve Web Service (SRW), to access data stored in the permanent repository.
- The delivery component delivers individual preserved intellectual entities via viewers that are supplied as part of the system (for example, a video streaming server) or as third-party applications. The system processes item delivery requests and checks the access rights defined for an item before delivering it to the user.

ExLibris Rosetta

Ex Libris Marketing Movies

Alma Next-generation library
file pid: FL23568

bX article recommender plain and simple
file pid: FL23569

Primo Central plain and simple
file pid: FL23570

Primo ScholarRank plain and simple
file pid: FL23571

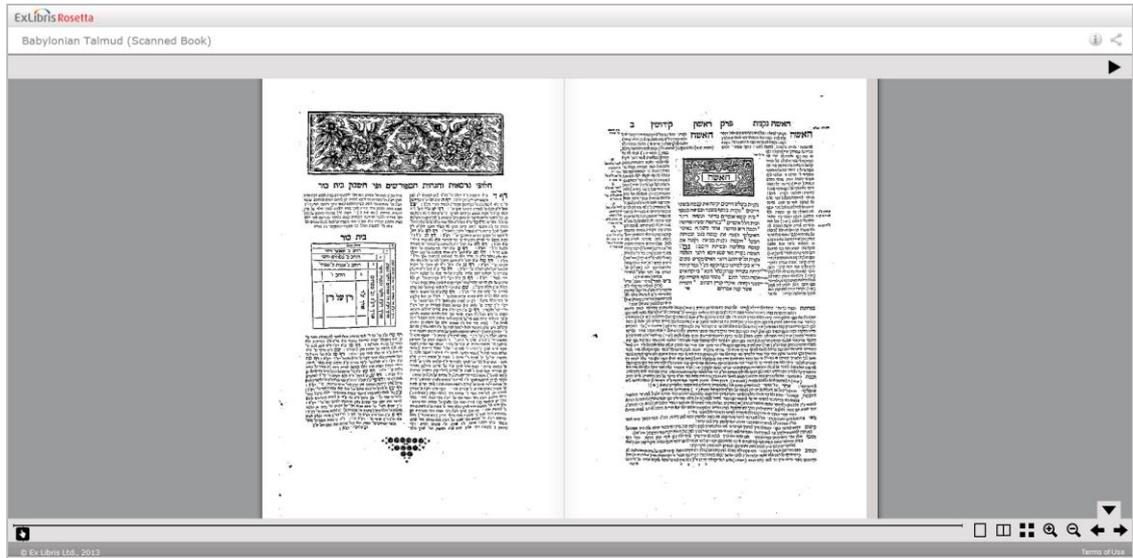
Rosetta plain and simple
file pid: FL23572

Yonsei University Library and Ex Libris
file pid: FL23573

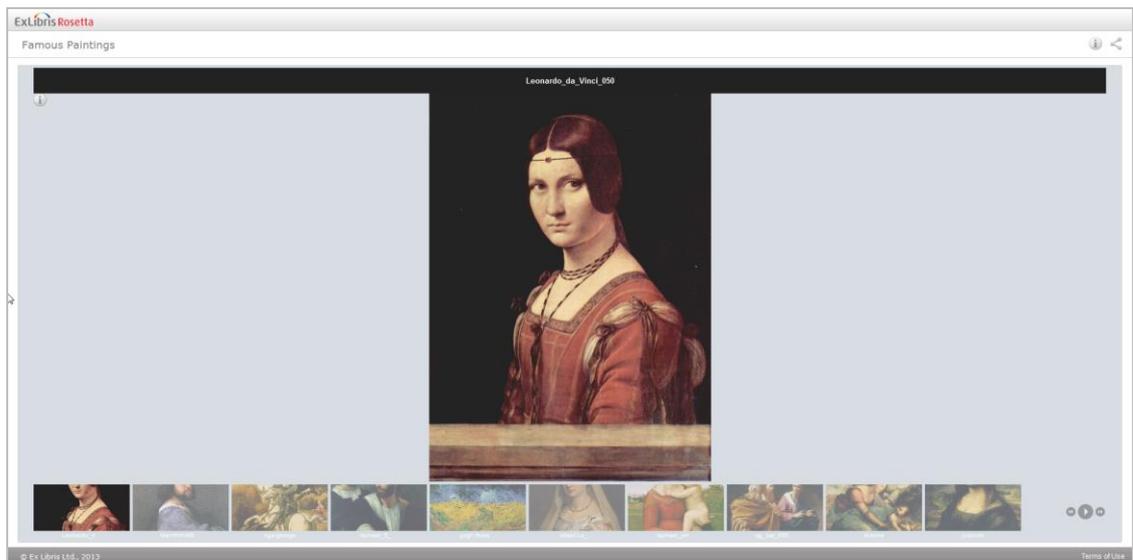
00:02 / 03:10

© Ex Libris Ltd., 2013 Terms of Use

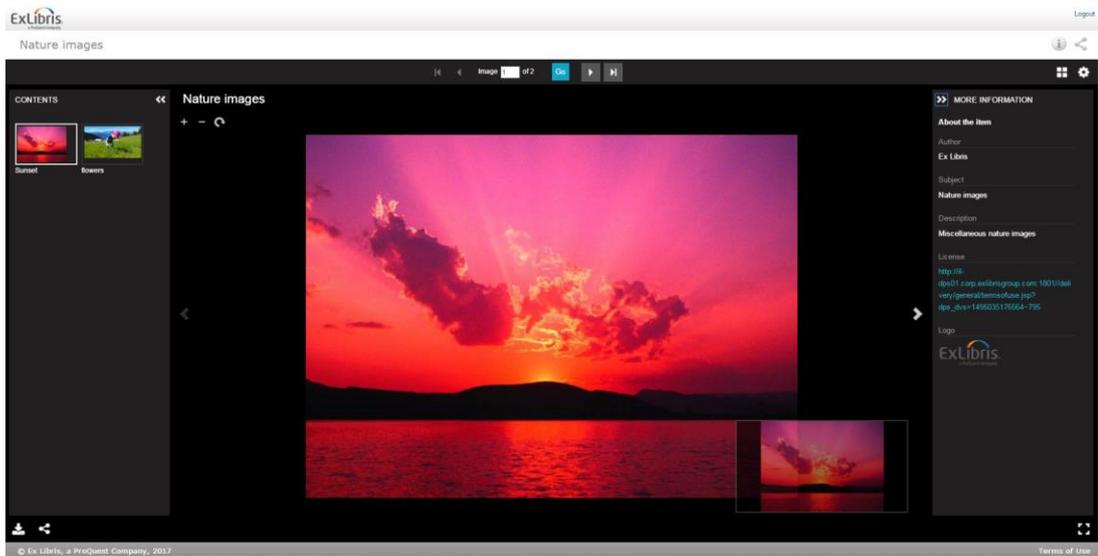
Delivery Module: Video Player



Delivery Module: IA Book Redaer



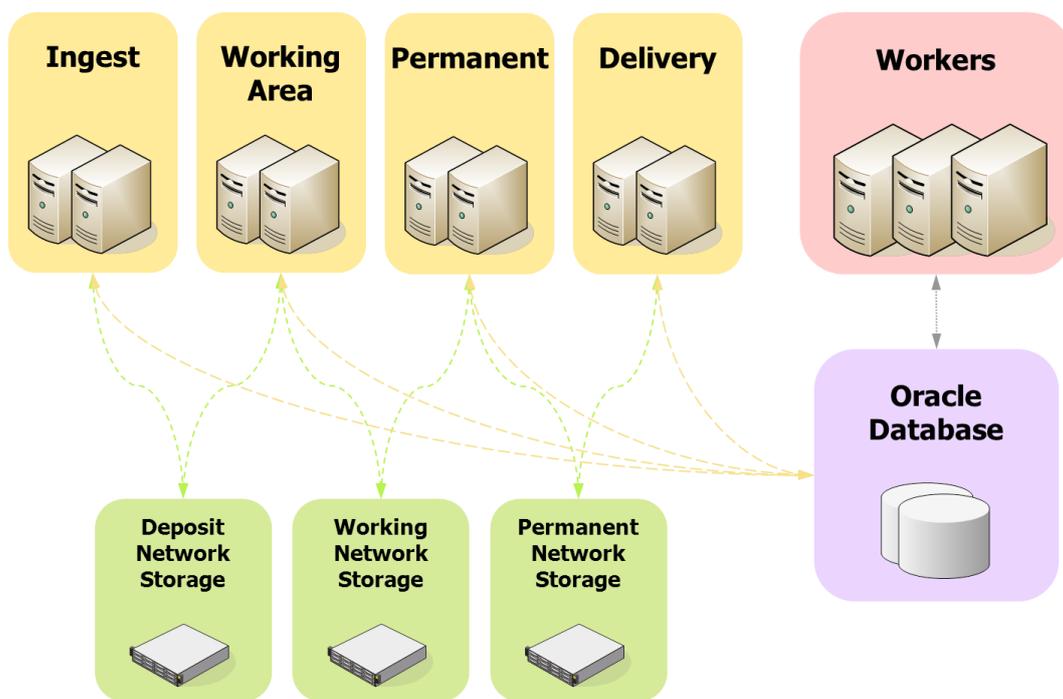
Delivery Module: Photo Album Viewer



IIIF Universal Viewer

Scalability

Ex Libris Rosetta provides a scalable infrastructure to address the ever-growing need to preserve and manage digital materials. This scalable solution can be implemented on a robust distributed architecture which allows the deposit module, working area, permanent repository, and database to be deployed on separate servers. Each module can be scaled up with additional of computers. In addition, Ex Libris Rosetta enables institutions to add dedicated servers, called *workers*, to perform specific tasks, such as virus and fixity checks. This flexibility allows an institution to start with a small hardware configuration and expand Rosetta to meet the needs of the institution's growing collection.



Scalability Diagram

Beyond the benefit of scalability, the architecture provides institutions with a system that offers redundancy and has no single point of failure.

In addition, Amazon S3 Plugins and other plugins for Tapes, HSMN, and HDFS can be created using the plugin module.

Openness

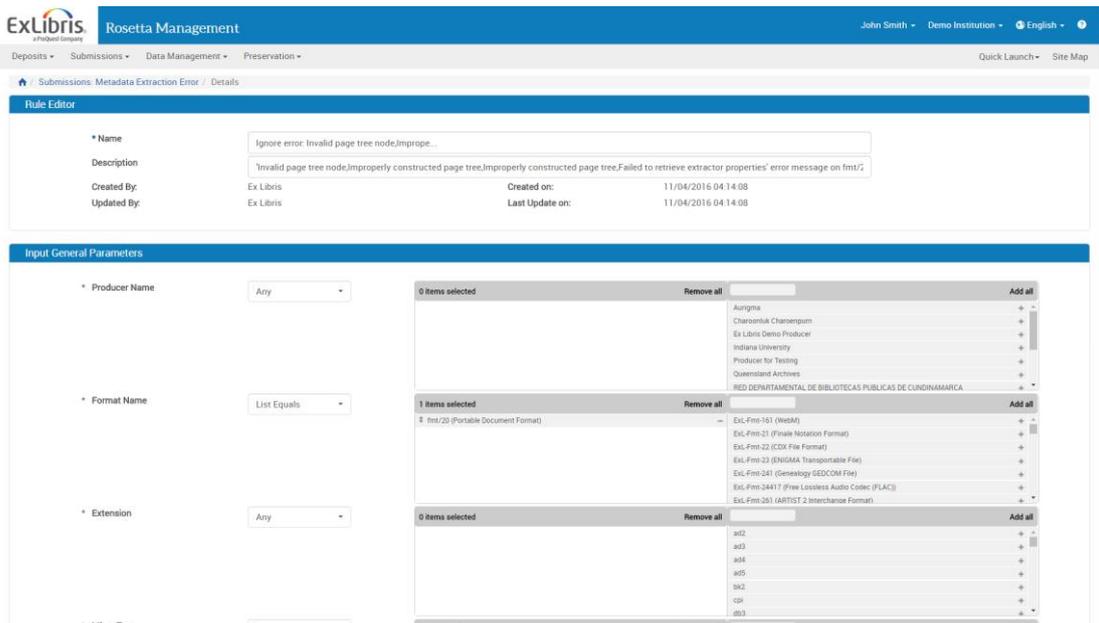
The Rosetta platform supports ingest, management, preservation, and delivery processes. Its open architecture combined with a rule-based workflow engine enables institutions to enhance the system with plug-in tools and applications in addition to those offered out-of-the-box. Such tools support the following tasks and activities:

- **Submission:** A set of APIs and a software development kit (SDK) enable institutions to create unique submission applications that support integration with existing institutional applications and can be seamlessly integrated with Rosetta.
- **Characterization:** External third-party tools such as JHOVE, DROID, and virus check applications are embedded in Rosetta to ensure that information in the system is viable and not damaged. New tools can be added as they emerge or evolve.
- **Enrichment:** Managed by a workflow engine, the rule-based ingest solution allows institutions to embed a wide variety of tools that add information to the AIPs stored in the system—for example, tools that convert file formats such as AVI to MP3. External metadata management systems (such as an ILS catalog or an archival system) can be additional sources of metadata enrichment.
- **Dissemination:** The dissemination SDK is a set of tools that enables Rosetta to be integrated easily with local systems for digital rights management, e-commerce, and discovery.

Business Logic-Driven Workflows

A business rules engine is used in the various Rosetta modules. For example, if the customer would like that specific materials coming from specific publisher will go through specific validation procedure, the customer can configure the system to recognize such materials and redirect them to the required validation. This mechanism is used in various places in Rosetta, starting from the ingest workflows through preservation mechanisms and finally the delivery module.

Each business rule contains input parameters and output parameters, where the input parameters are matched the system acts according to the defined process in the output parameters.



Metadata Extraction Error Rules

Rosetta Community-Driven Knowledge Base

One of the core features of Rosetta is the format library (a.k.a. community driven knowledge base), it's a mechanism to share and store information about formats, application, technical metadata extractors and risks. This KB is managed and updated by Rosetta users' community where Ex Libris provides the infrastructure that allows smooth and easy sharing.

7

The screenshot shows the 'List of Formats' page in the Rosetta Management system. The page includes a search bar and a table of 16 format entries. The table columns are: Associated/Disassociated, Name, Description, Version, Classification, Registry Type, and Registry ID. Each row has 'View' and 'Edit' links.

Associated/Disassociated	Name	Description	Version	Classification	Registry Type	Registry ID		
1 Disassociated	ExL-Fmt-21	Finale Notation Form...	-	Generic	EX Global	ExL-Fmt-21	View	Edit
2 Disassociated	ExL-Fmt-22	CDX File Format	-	Text (Unstructured...	EX Global	ExL-Fmt-22	View	Edit
3 Disassociated	ExL-Fmt-23	ENIGMA Transportable...	-	Generic	EX Global	ExL-Fmt-23	View	Edit
4 Disassociated	ExL-Fmt-41	Open Publication Str...	2.0 v1.0	eBook	EX Global	ExL-Fmt-41	View	Edit
5 Associated	ExL-Fmt-61	MPEG-4 Media File	-	Generic	PRONOM	fmt/199	View	Edit
6 Associated	ExL-Fmt-62	Microsoft Office Ope...	2007	Generic	PRONOM	fmt/189	View	Edit
7 Disassociated	ExL-Fmt-161	WebM	-	Video	EX Global	ExL-Fmt-161	View	Edit
8 Disassociated	ExL-Fmt-241	Genealogy GEDCOM Fil...	5.5	Text (Structured)	EX Global	ExL-Fmt-241	View	Edit
9 Disassociated	ExL-Fmt-261	ARTIST 2 Interchange...	2.1.2	eBook	EX Global	ExL-Fmt-261	View	Edit
10 Disassociated	ExL-Fmt-321	Webarchive	-	Webharvest	EX Global	ExL-Fmt-321	View	Edit
11 Associated	ExL-Fmt-24417	Free Lossless Audio ...	-	Audio (NLNZ)	PRONOM	fmt/279	View	Edit
12 Associated	fmt/1	Broadcast WAVE	0	Audio (AES)	PRONOM	fmt/1	View	Edit
13 Associated	fmt/2	Broadcast WAVE	1	Audio (AES)	PRONOM	fmt/2	View	Edit
14 Associated	fmt/3	Graphics Interchange...	1987a	Image (Mix)	PRONOM	fmt/3	View	Edit
15 Associated	fmt/4	Graphics Interchange...	1989a	Image (Mix)	PRONOM	fmt/4	View	Edit
16 Associated	fmt/5	Audio/Video Interlea...	-	Video	PRONOM	fmt/5	View	Edit

Rosetta's Format Library

Based on the information shared between Rosetta users, each customer can create preservation plans and preservation actions and to get the community feedback prior to any large impact actions.

Find Out More

For information about Ex Libris Rosetta, related events, and Webinar opportunities, go to the [Ex Libris Web site](#). Full Rosetta Documentation is available in the [Ex Libris Knowledge Center](#). Product APIs and plug-ins documentation are available from the [Ex Libris Developers Network](#).