



Rosetta 5.2 Highlights

March 2017

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. This information is provided AS IS. Unless otherwise agreed, Ex Libris shall not be liable for any damages for use of this document, including, without limitation, consequential, punitive, indirect or direct damages.

Any references in this document to third-party material (including third-party Web sites) are provided for convenience only and do not in any manner serve as an endorsement of that third-party material or those Web sites. The third-party materials are not part of the materials for this Ex Libris product and Ex Libris has no liability for such materials.

TRADEMARKS

"Ex Libris," the Ex Libris bridge, Primo, Aleph, Alephino, Voyager, SFX, MetaLib, Verde, DigiTool, Preservation, URM, Voyager, ENCompass, Endeavor eZConnect, WebVoyage, Citation Server, LinkFinder and LinkFinder Plus, and other marks are trademarks or registered trademarks of Ex Libris Ltd. or its affiliates.

The absence of a name or logo in this list does not constitute a waiver of any and all intellectual property rights that Ex Libris Ltd. or its affiliates have established in any of its products, features, or service names or logos.

Trademarks of various third-party products, which may include the following, are referenced in this documentation. Ex Libris does not claim any rights in these trademarks. Use of these marks does not imply endorsement by Ex Libris of these third-party products, or endorsement by these third parties of Ex Libris products.

Oracle is a registered trademark of Oracle Corporation.

UNIX is a registered trademark in the United States and other countries, licensed exclusively through X/Open Company Ltd.

Microsoft, the Microsoft logo, MS, MS-DOS, Microsoft PowerPoint, Visual Basic, Visual C++, Win32,

Microsoft Windows, the Windows logo, Microsoft Notepad, Microsoft Windows Explorer, Microsoft Internet Explorer, and Windows NT are registered trademarks and ActiveX is a trademark of the Microsoft Corporation in the United States and/or other countries.

Unicode and the Unicode logo are registered trademarks of Unicode, Inc.

Google is a registered trademark of Google, Inc.

Copyright Ex Libris Limited, 2017. All rights reserved.

Document released: March 2017

Author: Adi Alter, Rosetta Product Manager

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

Table of Contents

Contents

Infrastructure	4
1.1 SAML/Local Authentication	4
1.2 Cleanup Job	4
User Interface	6
Deposit and SIP Processing	7
3.1 BagIt Deposit	7
3.2 Improved TA Workbench and TA Error Visibility	7
3.3 OAI-PMH Test UI	8
3.4 Submission Job Reports	9
Data Management	10
4.1 Alma-Rosetta Digitization Flows	10
4.1.1 <i>Alma Submission Application</i>	11
4.1.2 <i>'Create Alma Digital Inventory' Repository Task</i>	11
4.1.3 <i>'Delete Alma Digital Inventory' Repository Task</i>	11
4.2 Delete Derivative Copy Representation Task	11
Preservation	12
5.1 Retaining Timestamps for Digital Content Uploaded to Rosetta	12
5.2 Tracking Changes to Access Rights Policies	12
5.3 Storage Migration	13
Publishing and Delivery	14
6.1 Search and SRU Improvements	14
6.1.1 <i>Exact Search for Identifiers</i>	14
6.1.2 <i>"Is Empty" Support for SRU</i>	14
6.2 Fulltext Viewer Preprocessor (VPP)	14
APIs	15
7.1 Access Rights Check Service	15
7.2 APIs Updates	15

1

Infrastructure

1.1 SAML/Local Authentication

Rosetta now supports authenticating users without the need to use a PDS server. Users can either be authenticated locally by Rosetta or can be authenticated by an external system using SAML (Security Assertion Markup Language) 2.0, an XML-based, open-standard data format for exchanging authentication and authorization data between parties.

SAML enables Rosetta to exchange authentication and authorization information with your institutional identity provider (IdP), allowing a single sign on for the institution's users. A user who will sign in or out of an external system will be automatically signed in or out of Rosetta, or vice versa.

1.2 Cleanup Job

A new cleanup job has been added to Rosetta's system jobs, deleting data older than a configurable number of days. The following areas are covered by this new job:

- Staff work area Directories
- Delivery Cache
- Update Metadata Job Directories
- Add Derivative Representation Job Directories
- Old Deposit Jobs
- Preservation Alternative Plans Directories

Job Parameters ▼

<input type="checkbox"/> Clean staff_work_area Directories	Older than days:	<input type="text" value="30"/>
<input type="checkbox"/> Clean Delivery Cache	Older than days:	<input type="text" value="90"/>
<input type="checkbox"/> Clean Update Metadata Job Directories	Older than days:	<input type="text" value="30"/>
<input type="checkbox"/> Clean Add Derivative Rep Job Directories	Older than days:	<input type="text" value="30"/>
<input type="checkbox"/> Clean Old Deposit Jobs	Older than days:	<input type="text" value="183"/>
<input type="checkbox"/> Clean Preservation Alternative Plans Directories	Older than days:	<input type="text" value="90"/>
<input type="checkbox"/> Clean Finished SIPs	Older than days:	<input type="text" value="183"/>

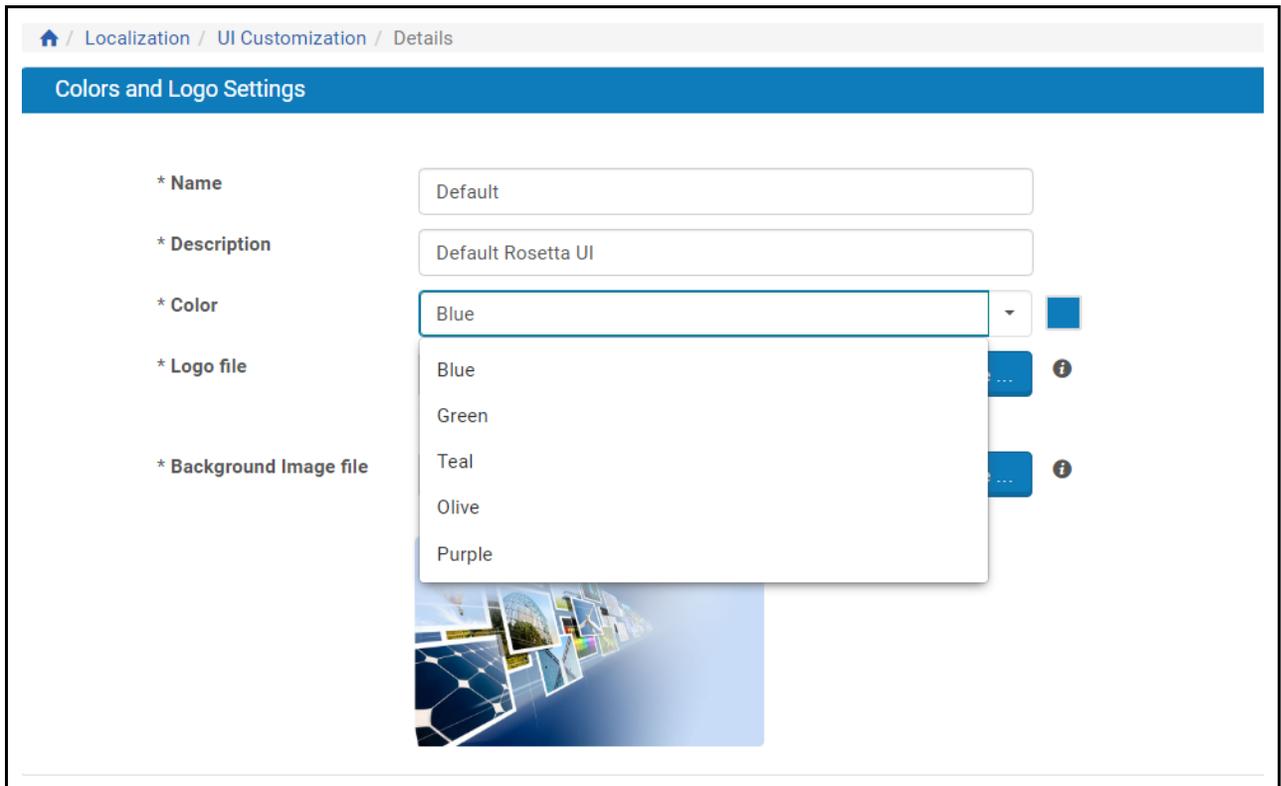
Cleanup Jobs Configuration

2

User Interface

We continued the work started in version 5.1 by further enhancing Rosetta’s UI and introducing multiple improvements throughout the system.

As part of this work, Rosetta now supports once more UI color customization. It is possible to choose a color scheme from the following five colors: purple, blue, teal, green and olive. For more information about these colors, their meaning, and their variance, refer to the Colors section available on [ProQuest’s UX Framework site](#).



UI Customization

3

Deposit and SIP Processing

3.1 BagIt Deposit

It is now possible to upload SIPs into Rosetta using BagIt, a hierarchical file packaging format designed to support disk-based storage and network transfer of arbitrary digital content.

BagIt bags are ideal for digital content normally kept as a collection of files. They are also well-suited to the export, for archival purposes, of content normally kept in database structures that receiving parties are unlikely to support. BagIt is easy to implement and is cross-platform, therefore making it widely adopted by many digital libraries.

The new BagIt Content Structure Converter enables uploading content in the BagIt format and mapping BagIt tags to Rosetta's metadata. Rosetta extracts the bag, deposits the files composing it into Rosetta, and creates an intellectual entity out of them.

The screenshot displays the 'Content Structure Details' interface. At the top, there is a blue header with the title and a dropdown arrow. Below the header, there are two input fields: '* Name' with the value 'BagIt' and '* Status' with a dropdown menu set to 'Active'. A checkbox labeled 'Store Tags as Source Metadata' is checked. The main part of the interface is a table with four columns: 'Tag File', 'Tag', 'Property', and an action column. The table contains four rows of data. Below the table, there are three input fields: 'Tag File', 'Tag', and '* Property', with a 'Create' button to the right. A search icon is located below the '* Property' field. At the bottom right, there are 'Cancel' and 'Save' buttons.

	Tag File	Tag	Property	
1	bag-info.txt	External-Identifier	SIP - Identifier (DC)	Remove
2	bag-info.txt	Bagging-Date	SIP - Date (DC)	Remove
3	bag-info.txt	External-Description	SIP - Description (DC)	Remove
4	bag-info.txt	Source-Organization	SIP - Publisher (DC)	Remove

BagIt Content Structure Converter

3.2 Improved TA Workbench and TA Error Visibility

The columns of the different tabs on the Technical Issues page have been standardized, ensuring consistent user experience across tabs.

In addition, a new **View Errors** action has been added to the relevant tabs. This action displays the error details of the relevant event, containing the failure reason. As part of this effort, the CSV Content Structure error logging was enhanced to include the problematic IE/REP information, thus enabling easier troubleshooting of the submission error.

The screenshot shows the ExLibris Rosetta Management interface. A modal window is open, displaying the following error message:

```
Found Multiple Derivative copy
representations with Representation
Entity Type = pdf and Representation
Code = low (dc:title 'Gutenberg Bible',
dc:identifier '123456')
```

The modal also includes a 'Close' button. In the background, a table of submissions is visible with columns for SIP ID, Deposit ID, Producer, and various actions like 'View Errors', 'Work On', 'Update', 'Assign to', and 'Unassign'.

3.3 OAI-PMH Test UI

The OAI-PMH Harvester allows users to load objects into Rosetta directly from an OAI-PMH digital repository. It is now possible to test the OAI-PMH Harvester job and check its connectability, record transformation, and match per selected job configuration, thus ensuring successful execution of the job once it is scheduled to run a full harvest.

The screenshot shows the OAI-PMH Test Area interface. It includes the following elements:

- Record:** A dropdown menu set to 'First Record' and a 'Test' button.
- Status:** A green indicator showing 'Success'.
- Source Record:** A text area containing XML metadata:


```
<xml-fragment xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
<ns:header xmlns:ns="http://www.openarchives.org/OAI/2.0/">
<ns:identifier>oai:coa1.loc.gov:loc.gdc/gckb.001</ns:identifier>
```
- Transformed Record:** A text area containing XML metadata:


```
<?xml version="1.0" encoding="UTF-8"?>
<dc:record xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:mods="http://www.loc.gov/mods/v3" xmlns:dnx="http://www.exlibrisgroup.com/dps/dnx"
xmlns:mets="http://www.loc.gov/METS/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:dcterms="http://purl.org/dc/terms/"
```
- Match:** A text field showing 'No match found'.
- Buttons:** 'Back', 'Refresh', 'Apply', and 'Run Now' at the bottom.

OAI-PMH Test Area

3.4 Submission Job Reports

The following two new reports were added to the Submission Reports section, enabling better tracking of submission jobs' status:

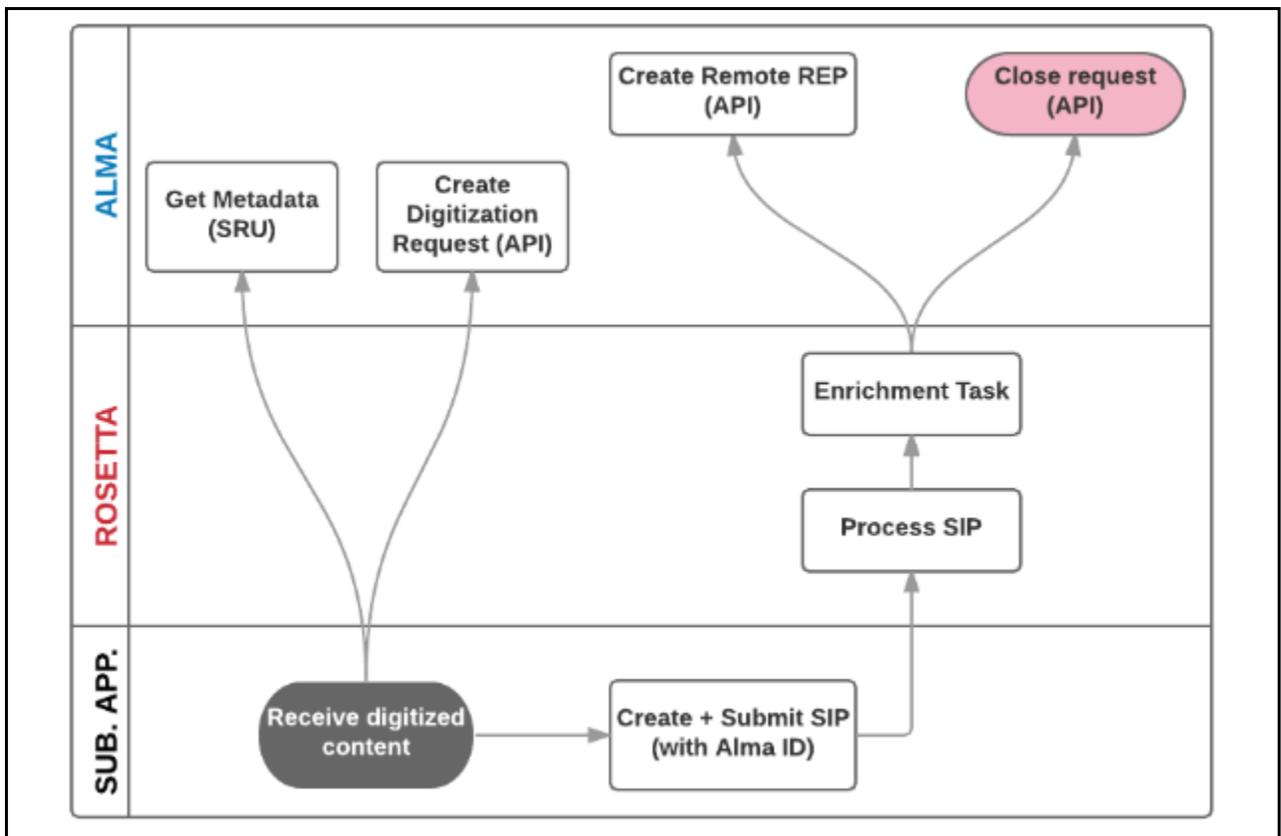
- Submission Job Details by Job Name and Execution Date
- Submission Job Details by Source Directory

4

Data Management

4.1 Alma-Rosetta Digitization Flows

The new Alma-Rosetta integrated digitization flow provides a submission application that creates Rosetta SIPs for digitized content and a Rosetta repository task that uses Alma APIs to create remote digital inventory for the ingested content, as can be seen in the following diagram:



4.1.1 Alma Submission Application

The Alma Submission Application is used for submitting to Rosetta digitized content related to an Alma bibliographic record (and, optionally, a fulfilment of a related digitization request). It retrieves descriptive metadata from Alma using SRU, creates a Rosetta METS file, and submits Rosetta SIPs for a set of files.

The application is based on the Rosetta SDK and Alma bibliographic APIs. The application code is available on GitHub and it can be used as is, or modified to fit specific needs and workflows.

4.1.2 'Create Alma Digital Inventory' Repository Task

This new repository task creates digital inventory in Alma using information coming from IEs created by the Alma submission application.

A successful **Create Alma Inventory** task results in the following:

- A new remote representation in Alma.
- If provided, the Alma request ID is added as a representation property, removed from the IE, and then the request is closed (if it is in the appropriate stage).

4.1.3 'Delete Alma Digital Inventory' Repository Task

The delete Alma inventory task can be manually run to delete digital inventory created incorrectly or to remove digital inventory before deleting the IE.

4.2 Delete Derivative Copy Representation Task

Derivative copy representations can now be deleted using a Rosetta job. This is especially helpful when you need to delete such copies in bulk.

5

Preservation

5.1. Retaining Timestamps for Digital Content Uploaded to Rosetta

Original file timestamps are now preserved for storage that uses the copy method for all kinds of submission formats, thus ensuring that original date characteristics are not modified when moving content into Rosetta.

5.2. Tracking Changes to Access Rights Policies

Access Rights, Access Rights Exceptions, and Retention Policies now have a **History** action, enabling users to view and revert to previous MD versions.

	Mid	Description	Metadata Type	Format	Created by	Creation Date		Delete	
1	1	Limited ac...	accessrights	policy	SYSTEM	3/21/17	Edit	Delete	-
2	2	Embargoe...	accessrights	policy	SYSTEM	3/21/17	Edit	Delete	History
3	3	No restrict...	accessrights	policy	SYSTEM	3/21/17	Edit	Delete	-
4	4	Accessible...	accessrights	policy	SYSTEM	3/21/17	Edit	Delete	-
5	5	From withi...	accessrights	policy	SYSTEM	3/21/17	Edit	Delete	-
6	1164	AR selleni...	accessrights	policy	admin1	3/21/17	Edit	Delete	-
7	1167	AR selleni...	accessrights	policy	admin1	3/21/17	Edit	Delete	History
8	1189	Delivery Re...	accessrights	policy	admin1	3/21/17	Edit	Delete	-
9	1195	Deny All	accessrights	policy	admin1	3/21/17	Edit	Delete	History
10	1330	Not Allowe...	accessrights	policy	admin1	3/21/17	Edit	Delete	History

1 - 10 of 21 Records

1 2 3 > >>

Access Rights Policies Management

MID	2	Metadata Type	policy:accessrights	Description	Embargoed for 5 Years
Created by	SYSTEM	Creation Date	21/03/2017 19:32:47		
Updated by	admin1	Update Date	22/03/2017 00:40:49		

22/03/2017 00:40:48 by John Smith		Revert
Policy Information		>

22/03/2017 00:40:14 by John Smith		Revert
Policy Information		>
Copyright Template	-	
Description	Embargoed for 5 Years	
Restricted Message	English:	
Group1	Moving Wall after DNX:objectCharacteristics.creationDate 5 Embargoed for 5 Years	
Group2	User Group equal STAFF Embargoed for 5 Years	

22/03/2017 00:39:41 by John Smith		Revert
Policy Information		>

Access Rights Policies History Management

5.3. Storage Migration

Rosetta supports now migrating a set of files from one permanent storage location to another. The migration is based on recommitting files to the permanent storage. A fixity check confirms that the source and target files are identical. This is helpful, for example, when there is a need to move some of the objects stored on one storage location to another, while still leaving other objects on the original storage.

6

Publishing and Delivery

6.1 Search and SRU Improvements

6.1.1 Exact Search for Identifiers

Exact search is now supported for identifiers both via the UI using the `equals` operator and via SRU using the `==` operator.

6.1.2 "Is Empty" Support for SRU

The `Is Empty` operator is now supported in SRU, by using the `=="` operator.

6.2 Fulltext Viewer Preprocessor (VPP)

A new Viewer Preprocessor (VPP) that converts all PDF files in a representation into a single HTML file was added to Rosetta. This VPP should be used with the StreamGate viewer to provide a fulltext representation to a harvester (for example, Primo).

7

APIs

7.1 Access Rights Check Service

Rosetta now offers a service to check permissions on an object without actually delivering it. This is done by accessing the delivery URL using a HEAD HTTP request. Rosetta will return a 401 or 403 HTTP status code if access is denied and it may also return a denied note. Such a request does not count as a delivery request.

Note that this service is now used by Alma to indicate in the service tab (ViewIt) to a patron that access to a remote representation is denied by the remote repository.

7.2 APIs Updates

- The getRipStatus web service was enhanced to return a **Finished** Status for successfully added/updated representations.
- Required roles for executing web services have been updated thus providing improved web services security.
- Web services Javadoc documentation has been improved