



Rosetta 5.1 Highlights

September 2016

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, 2016. All rights reserved.

Document released: September 2016

Author: Adi Alter, Rosetta Product Manager

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

Table of Contents

Contents

Infrastructure	5
1.1 New Third-Party Software Versions	5
User Interface	6
Deposit and SIP Processing	9
3.1 SIP Processing Prioritization	9
3.2 Submission Job History	9
3.3 Validation Stack	9
3.3.1 <i>SHA-256 Fixity Check</i>	9
3.3.2 <i>Virus Check Exit Codes</i>	10
3.3.3 <i>Rules for Automatic Handling of Virus Check Errors</i>	10
3.3.4 <i>Technical Analysts Rules Export</i>	11
Data Management	12
4.1 XML Editor	12
4.2 Add Derivative Copy Representation Job	14
4.3 Search	14
4.3.1 <i>'Not' Operator Support</i>	14
4.3.2 <i>Smart-Search for all Controlled-Vocabulary Fields</i>	14
4.4 Collection Management	14
4.4.1 <i>Collection Ordering</i>	14
4.4.2 <i>Disassociating IEs from Collections</i>	14
4.4.3 <i>Collection History Tab</i>	15
4.4.4 <i>Assign Collection by DC Improvements</i>	15
Preservation	16
5.1. File Comparison Plug-in	16
5.2. Create a Preservation Plan for More than One Source Format	16
5.3. Dead Reference Identification	17
5.4. Events Documentation	17
Publishing and Delivery	18
6.1 Enrich Publishing with Derivative Copies	18

6.2	Representation Viewer	19
6.3	Mobile Support	19
6.4	Setting Viewer Parameters	19
6.5	Allow External Viewers to be Called without Rosetta Delivery	Page19

APIs		20
-------------	--	-----------

1

Infrastructure

1.1 New Third-Party Software Versions

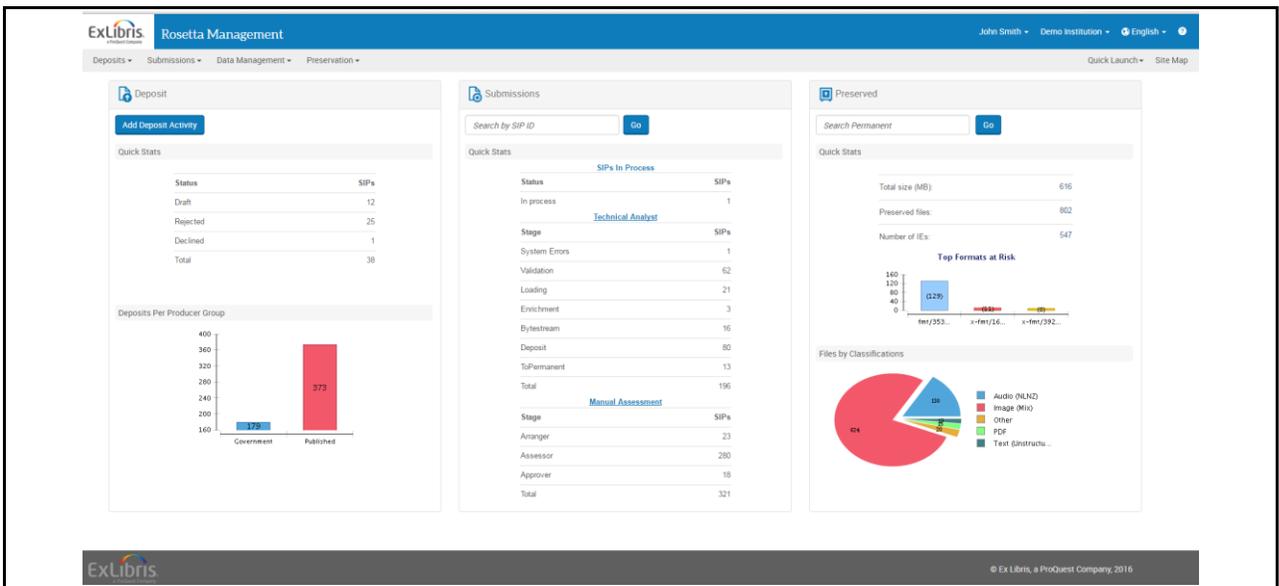
Rosetta now includes the followings new versions of third-party software:

- Tomcat 7.0.70
- PDS 2.1.7
- Aurigma Uploader 8.5.55
- FlexPaper 2.4.1
- JWPlayer 7

2

User Interface

This version introduces a new user interface with a new design of multiple widgets throughout the system – tables, forms, menus, breadcrumbs, etc. Some screenshots are provided below:



Main Dashboard

The screenshot shows the 'List of Formats' table with the following data:

Associated/Disassociated	Name	Description	Version	Classification	Registry_Type	Registry_ID	View	Edit
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-28	ZIV0k4desc.SuffixZoSj...	-	Audio (AES)	PRONOM	ExL-Fmt-28regIDSuffix	View	Edit
5 Disassociated	ExL-Fmt-41	Open Publication Str...	2.0 v1.0	eBook	EX Global	ExL-Fmt-41	View	Edit
6 Associated	ExL-Fmt-61	MPEG-4 Media File	-	Generic	PRONOM	fmt/199	View	Edit
7 Associated	ExL-Fmt-62	Microsoft Office Ope...	2007	Generic	PRONOM	fmt/189	View	Edit
8 Disassociated	ExL-Fmt-161	WebM	-	Video	EX Global	ExL-Fmt-161	View	Edit
9 Disassociated	ExL-Fmt-241	Genealogy GEDCOM Fil...	5.5	Text (Structured)	EX Global	ExL-Fmt-241	View	Edit
10 Disassociated	ExL-Fmt-261	ARTIST 2 Interchange...	2.1.2	eBook	EX Global	ExL-Fmt-261	View	Edit

Tables

ExLibris
a ProQuest Company

Rosetta Management

Deposits ▾ Submissions ▾ Data Management ▾ Preservation ▾

Producers and Agents	Deposit Arrangements	Advanced Tools
Producers	Material Flows	Run Reports
Producer Profiles	Content Structure	Schedule Reports
Producer Groups	Metadata Form	Delivery XSL Files
1 st Time Registration Reasons	Submission Format	Delivery Copyrights Statements
1 st Time Registration Rules	Metadata Profiles	Email Configuration
Policies	Assertion of Copyrights	OAI Harvester Transformation
Access Right Policies	Copyrights Boilerplate	Terms of Use Configuration Files
Access Rights Exceptions	Access Right Copyrights	
Retention Policies	Material Type	
	CSV Templates	
	Jobs	
	Producers Reports Job	
	Submission Job	
	OAI Harvester Job	

Actions Menu

ExLibris
a ProQuest Company

Rosetta Management

Deposits ▾ Submissions ▾ Data Management ▾ Preservation ▾

[Home](#) / [Preservation: Formats](#) / [Details](#)

Breadcrumbs

Material Flow Definition

* Name:

Description/Instructions:

Status: Active Status Date: 30/08/2016

Sampling Rate (%):

* Internal: No Yes

Material type:

Assertion of Copyrights:

Metadata Profile:

Technical Definitions

* Select content structure:

* Select submission format:

Automatically extract compressed files

Forms

Home

1
2
3
4
5

Descriptive Information

* Creator:

* Title:

* Description:

* Content Type:

* Creation date:

Publisher's name:

Notes:

Wizard

3

Deposit and SIP Processing

3.1 SIP Processing Prioritization

SIP submission and SIP processing queues now allow fast-tracking certain SIPs, controlled by a new **Priority** parameter defined on the SIP Processing Configuration page. SIPs with a higher priority are handled before SIPs with a lower one:

The screenshot shows the 'Submissions: SIP Processing Configuration / Details' page. The 'General Information' section includes fields for Name, Description, Priority, and Validation Stack Routine. The Priority dropdown menu is open, showing options: High, Normal, and Low. The Validation Stack Routine dropdown menu is also open, showing the option: Validation Stack - Full.

General Information	
* Name	General Digital Material
Description	All SIPs using generic or staff-mediated MF with (material type) t
Priority	Normal
Validation Stages - Pre-approval Automatic Stage	
Validation Stack Routine	Validation Stack - Full

3.2 Submission Job History

The jobs listed on the Submission Job and Job History pages were enhanced with additional data, enabling better analysis of the status of the submission jobs executed in Rosetta.

3.3 Validation Stack

3.3.1 SHA-256 Fixity Check

The SHA-256 fixity algorithm was added to the list of available fixity algorithms. SHA-256 is a [SHA-2](#) (Secure Hash Algorithm 2) cryptographic hash function computed with 32-bit words. It

is recommended to replace the SHA-1 algorithm with SHA-256. Due to collision attacks the use of this algorithm is no longer recommended for applications that depend on collision resistance, such as digital signatures. These attacks have not been successfully extended to SHA-2.

3.3.2 Virus Check Exit Codes

A new VirusCheckPluginV2 interface was added. This plug-in returns an integer for the virus scan result, as follows:

- 0 – pass
- 1 – failed
- >1 - undetermined

Technical Analysts can decide to ignore SIPs that fail TA Validation due to an undetermined virus check result.

The screenshot shows a web interface for file management. At the top, there are tabs for 'Content List', 'Metadata', 'Notes', and 'History'. Below the tabs is a 'Filter' dropdown set to 'All'. A table displays file records. The first record is 'Sunset.jpg' with PID 'FL3154' and Status '1 Tasks failed'. To the right of the table are buttons for 'Download', 'Replace', 'Recheck', and 'More...'. Below the table, a detailed view for the selected file shows the following information: Fixity Check: Pass, Virus Check: Undetermined - File not checked as it's over 2GBs (vs_Error.32868), File Format: Pass, Technical MD Extract: Pass, and Risk Analysis: Pass. At the bottom right of the interface, it says '1 - 1 of 1 Records'.

3.3.3 Rules for Automatic Handling of Virus Check Errors

Virus check errors can now be automatically handled by validation stack rules enabling Technical Analysts to create rules to ignore subsequent files with similar errors, as can be seen below. These rules are indexed and searchable as **Virus Check Error Ignore Criteria**.

Deposits ▾ Submissions ▾ Data Management ▾ Preservation ▾

🏠 / Submissions: Manage Issues in SIP Processing / Details

The following virus check error will be ignored:

Remember my decision concerning error message

Reason

Optional Note

These rules can be accessed in the new Virus Check Errors page that can be found under the **Submissions** menu:

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

🏠 / Submissions: Virus Check Error

Filter: All Find: in: All Go

[Add Rule](#) [Export Rules](#) 1 - 1 of 1 Records

Order	Name	Description	Creation Date	Modification Date				
1	<input checked="" type="checkbox"/>	Ignore virus check failure of big files	Error message will be ignored.	01/09/2016 18:56:39	01/09/2016 18:58:15	Update	Duplicate	Delete

1 - 1 of 1 Records

[Back](#)

3.3.4 Technical Analysts Rules Export

Technical Analysts can now export rules to a CSV file to share them with other institutions, enabling, for instance, institutions and consortia to share rules between each other.

4

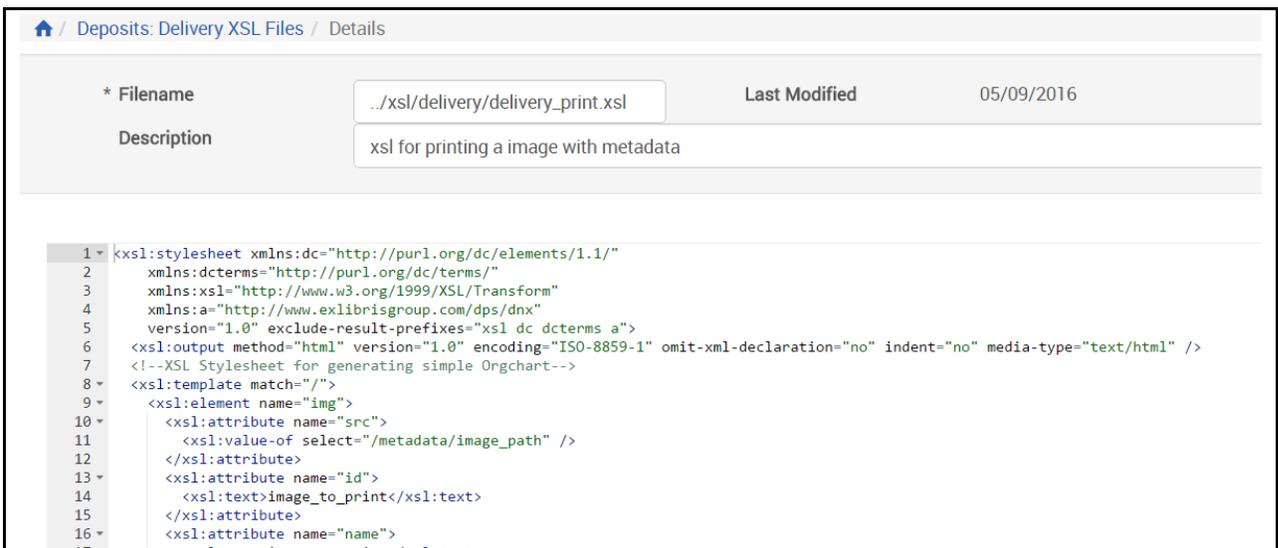
Data Management

4.1 XML Editor

A new XML editor was implemented, providing graphic tools such as indentation, tag autocomplete, and color-coding for editing the following files:

- Configuration files
- Source Metadata
- Logical Structmaps

The following is an example of a file edited using this editor:



The screenshot shows a web-based XML editor interface. At the top, there is a breadcrumb navigation: [/ Deposits: Delivery XSL Files / Details](#). Below this is a table with two columns: 'Filename' and 'Last Modified'. The filename is `../xsl/delivery/delivery_print.xml` and the last modified date is `05/09/2016`. A 'Description' field contains the text `xsl for printing a image with metadata`. Below the table is a code editor showing XSL code with line numbers 1 through 17. The code is as follows:

```
1 <xsl:stylesheet xmlns:dc="http://purl.org/dc/elements/1.1/"
2   xmlns:dcterms="http://purl.org/dc/terms/"
3   xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
4   xmlns:a="http://www.exlibrisgroup.com/dps/dnx"
5   version="1.0" exclude-result-prefixes="xsl dc dcterms a">
6   <xsl:output method="html" version="1.0" encoding="ISO-8859-1" omit-xml-declaration="no" indent="no" media-type="text/html" />
7   <!--XSL Stylesheet for generating simple Orgchart-->
8   <xsl:template match="/">
9     <xsl:element name="img">
10      <xsl:attribute name="src">
11        <xsl:value-of select="/metadata/image_path" />
12      </xsl:attribute>
13      <xsl:attribute name="id">
14        <xsl:text>image_to_print</xsl:text>
15      </xsl:attribute>
16      <xsl:attribute name="name">
```

In addition, Structmaps and Source MDs provided with an XSD file can be edited with the Editor's Grid mode, providing a contextual, schema-driven menus for adding new elements, sub-elements, and attributes. It is possible to upload any XSD file via Rosetta's Administration page and Rosetta will automatically generate the Grid more. The following are examples of the editor in Text and Grid mode for editing a logical structmap:

Home / Data Management: Search for Objects / Details

MID	Metadata Type	mets_section:structMap	Description	desc
Created by	Creation Date	8/18/16		
Updated by	Update Date	8/18/16		

Edit Select Insert Text Grid

```

1 <mets:structMap xmlns:mets="http://www.loc.gov/METS/"
2   ID="" TYPE="LOGICAL">
3   <mets:div LABEL="Preservation Master">
4     <mets:div LABEL="Table of Contents">
5       <mets:div LABEL="door.jpg" TYPE="FILE">
6         <mets:fptr FILEID="FL1005" />
7       </mets:div>
8     </mets:div>
9   </mets:div>
10 </mets:structMap>

```

Add Subelement
 mets:div
 Add Attribute
 mets:ID
 mets:TYPE
 Add Nodes
 Add Top Element
 mets:div

Validate Cancel Save

Text Mode

Edit Select Insert Text Grid

mets:structMap

(x) mets:ID REP1355-2 (x) mets:TYPE LOGICAL

mets:div

(x) mets:LABEL Test Struct Map

mets:div

(x) mets:LABEL Table of Contents

mets:div

(x) mets:LABEL Chapter 1

mets:div

(x) mets:LABEL Page 1 (x) mets:TYPE FILE

mets:fptr

(x) mets:FILEID FL1356

Add Subelement
 mets:fptr
 mets:div
 Add Attribute
 mets:ID
 mets:ORDER
 mets:ORDERLABEL
 mets:LABEL
 mets:TYPE
 Add Nodes
 Add Element
 Add Attribute
 Add CDATA
 Add comment
 Add Top Element
 mets:div

Grid Mode

4.2 Add Derivative Copy Representation Job

Derivative copy representations can now be added using a Rosetta job. This is especially helpful when you need to add externally-generated derivative copies in bulk.

4.3 Search

4.3.1 'Not' Operator Support

The operators **Does not contain keywords** and **Does not contain phrase** were added to the text properties search.

4.3.2 Smart-Search for all Controlled-Vocabulary Fields

The search fields for Producer, Producer Agent Names, Producer Authoritative Name, and Producer ID are now available as smart-search/auto-complete fields.

4.4 Collection Management

4.4.1 Collection Ordering

Rosetta now supports changing the order that collections appear in Rosetta. You can change the order of collections in the following ways:

- Dragging-and-dropping a collection in Rosetta's Collection Management UI.
- Using Rosetta's Create and Update collection APIs
- Providing the order of the collection in a CSV file while depositing content

The defined order will be the order in which collections are displayed in the Collection Viewer and in Primo's Collection Discovery lobby.

4.4.2 Disassociating IEs from Collections

Rosetta does not allow deleting a collection that has associated IEs (or sub-collections with associated IEs). The Unassign Collection process was added that disassociates all IEs from the collection and its sub-collections, so that you can delete the collection from Rosetta.

4.4.3 Collection History Tab

A new **History** tab was added to the Collection Manager. This tab displays a list of the actions performed on the collection, for example:

The screenshot shows the 'Collection Management' interface. At the top, the breadcrumb is 'Collection Management'. Below it, there is a search bar and a collection icon. The collection details are as follows:

Collection Name	CollectionByRosettaId9fCH5gl	Collection	Van Gogh
Collection ID	309914	Description	
Modification Date	9/6/16	Creation Date	9/5/16
External System	-	Publish	No
Allow Navigation	Yes	External Id	-

There is a 'View Collection' button on the right. Below the details are tabs for 'Contents', 'Metadata', and 'History'. The 'History' tab is active, showing a table of events:

Event type	Description	Date	
1	Collection has been updated	Collection: Van Gogh has been updated (collection id:309914)	06/09/2016 01:15:58
2	IE has been added to collection	IE4087 has been added to collection:CollectionByRosettaId9fCH5gl (...)	05/09/2016 18:52:37
3	Collection has been created	Collection: Van Gogh has been created (collection id:309914)	05/09/2016 18:51:10

At the bottom right of the history table, it says '1 - 3 of 3 Events'.

4.4.4 Assign Collection by DC Improvements

The **Assign Collection by DC** task now supports assigning by Rosetta ID or by an external system ID. The task also now supports a **Delete on Assignment** option that can be used when the DC field exists solely for collection assignment. In addition, users can now define whether or not to create new collections, when assigning a collection by a path.

5

Preservation

5.1. File Comparison Plug-in

A new plug-in type, **File Comparison Plug-in** was added to Rosetta. This plug-in can be used as part of a preservation evaluation to evaluate the quality of the migration. It receives the original file and the migrated file and returns a value that determines its quality.

An example of such a plug-in that computes PSNR (peak signal-to-noise ratio) between original and alternative images during preservation action as a visual method of evaluation criteria was added to [Ex Libris' GitHub repository](#).

5.2. Create a Preservation Plan for More than One Source Format

A preservation plan can now be created for more than one format as long as they are from the same classification and share the same risk. The purpose of this development is to simplify the creation of such a plan and avoid the need to create a new plan for each format. The following is an example of creating such a plan:

The screenshot shows the Rosetta Management interface. At the top, there is a navigation bar with the ExLibris logo and the text "Rosetta Management". Below this, there are several tabs: "Deposits", "Submissions", "Data Management", and "Preservation". The "Preservation" tab is active. Below the tabs, there is a breadcrumb trail: "Preservation: View Global Risk Report / Details". The main content area shows a "Risk Type" of "Zero Applications". Below this, there is a "Filter" dropdown set to "Image (Mix)" and a "Find" search box. A table with 10 columns is displayed, showing two rows of data. The columns are: a checkbox, "Format Name", "Extension", "Classification", "Institution", "No. of IEs", "No. of Representations", "No. of Files", "Create Set", and "Preservation Plans". The first row has a checked checkbox, "fmt/353", "tif,tiff.ptif", "Image (Mix)", "Demo Institution", "104", "113", "130", "Create Set", and "Preservation Plans". The second row has a checked checkbox, "x-fmt/392", "jp2", "Image (Mix)", "Demo Institution", "6", "6", "6", "Create Set", and "Preservation Plans". At the bottom of the table, there are three buttons: "Back", "Create Set for Selected", and "Preservation Plans for Selected".

	Format Name	Extension	Classification	Institution	No. of IEs	No. of Representations	No. of Files	Create Set	Preservation Plans
1	fmt/353	tif,tiff.ptif	Image (Mix)	Demo Institution	104	113	130	Create Set	Preservation Plans
2	x-fmt/392	jp2	Image (Mix)	Demo Institution	6	6	6	Create Set	Preservation Plans

5.3. Dead Reference Identification

A new type of fixity check was added to Rosetta to ensure data integrity and that all files referenced in the database are present in the storage. This check does not run an actual fixity algorithm, but simply verifies that the file with the expected name exists in storage without actually accessing the file to determine its integrity.

5.4. Events Documentation

A table listing the details of all the events that are managed by Rosetta was added as an appendix to the Rosetta Configuration guide:

ID	Description	Preservation	Audit	Statistic
207	Process Automation framework result	N	N	N
208	Arranger - Decline SIP	N	Y	Y
209	Assessor - Forward SIP	N	N	N
210	Process Ended Execution	N	Y	N
211	Object has been locked	Y	N	N
212	Object has been rolled back	Y	Y	Y
213	Object has been committed	Y	Y	Y
214	Commit/Rollback error	N	Y	Y
215	SIP moved between stages	N	N	N
216	Started MD Validation Stage	N	N	N

6

Publishing and Delivery

6.1 Enrich Publishing with Derivative Copies

Users can now indicate the preservation type when adding publishing profiles and so publish derivative copy representation metadata, for example:

Home / Data Management / Manage Sets and Processes / Publishing Configuration / A

* Name Status Active ▾

Description

Preservation Type 0 items selected Remove all Add all

- Preservation Master
- Modified Master
- Derivative Copy

* Converter Type OaiCollectionConverterPlugin ▾ Include CMS

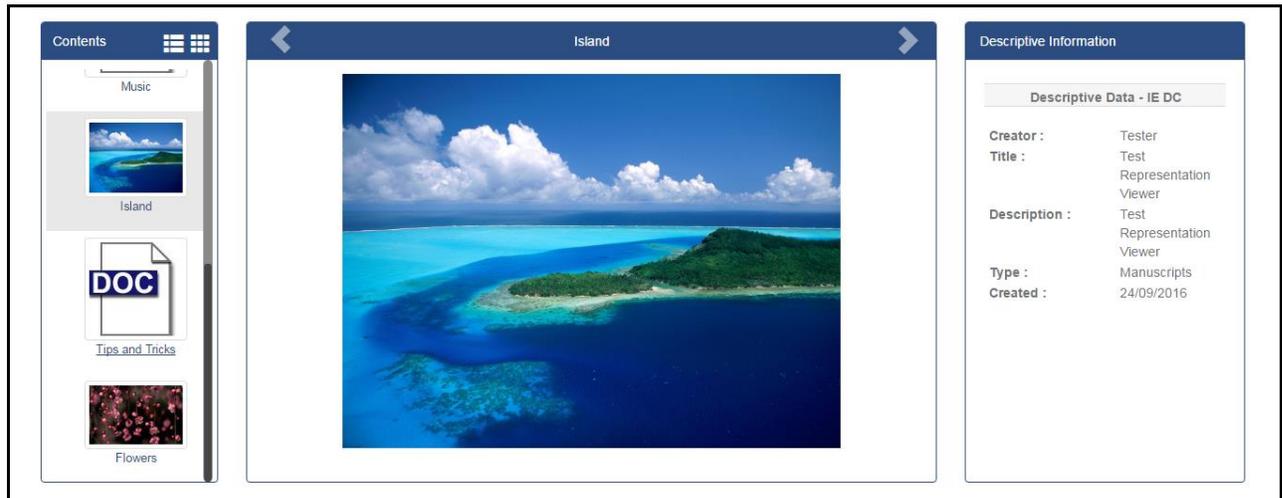
* Target Type NFSPublisherPlugin ▾ Include Access Rights

Cancel Next >

Similar UI improvements were applied to the **Export Intellectual Entities** task.

6.2 Representation Viewer

A new HTML5 viewer that accepts a single representation was added to Rosetta. This viewer can display all formats that are natively supported by the browser, such as images, audio, and video content, etc. For other formats, the viewer supports download. The following is an example of the viewer:



6.3 Mobile Support

Both the JWPlayer and the FlexPaper viewers were upgraded to new versions that supports HTML5, enabling the use of these viewers on devices that do not support Adobe® Flash®, such as mobile devices.

6.4 Setting Viewer Parameters

Any non-reserved query string parameter can now be appended to the delivery request and passed to the viewer. This can be done by appending the parameter directly on the request URL or by adding it via the delivery rules.

6.5 Allow External Viewers to be Called without Rosetta Delivery Page

It is possible now to open an external viewer without the Rosetta wrapper, that is, without the Rosetta header, footer, and the toolbar with the login/logout links.

7

APIs

The following new APIs were added:

- **getSIPStatusInfoByExternalId** – gets the SIP status through external ID and external system.
- **getSIPsStatusInfo** – returns a list of up to 1,000 **SipStatusInfo** objects.

The following APIs were updated:

- **UpdateMD** Web service's method now supports adding and updating a representation's logical structMap.
- **getIE** Web services can return now the file physical location, if required.
- Access Rights and Access Rights Exceptions web services support now also derivative copies.
- When access to a stream is denied by an access rights policy, a **401 Unauthorized status** code is returned.
- The **Locked IE** exception, which occurs when a Web service fails to run on an IE because the IE is locked, now includes a locking agent ID.