Rosetta Product Update

IGeLU, September 2016

Adi Alter | Rosetta Product Manager
Growing Number of Institutions Across all Regions

Serving more than 200 Institutions from 17 Countries
Installed Base Covering Various Institution Types

National Libraries & Institutions
- National Library Board (Singapore)
- TIB (German National Library of Science and Technology)
- Te Puna Mātauranga a Aotearoa (New Zealand)
- Slovenská Národná knižnica (Slovak National Library)

State Libraries
- BSB (Bayerische Staatsbibliothek)
- BVB (BibliotheksVerbund Bayern)
- SL (State Library of Queensland)
- OhioLINK

Academic Institutions
- Binghamton University
- University of Nebraska, Lincoln
- Memorial University of Newfoundland
- ETH-Bibliothek (Swiss Federal Institute of Technology, Zurich)
- Shanghai Library

Archives, Governmental, Public Libraries & Cultural Heritage
- Center for Jewish History
- Archiv Niederösterreich
Main Focus Areas Since Last IGeLU

1. Enhance Core Functionalities
2. Increase Synergy with other Ex Libris Products
3. Streamline the Implementation Process
Enhance Core Functionalities

- Digital Asset Management
  - Content Delivery
  - Data Management
  - Collection Management

- Preservation
  - Community-driven Knowledge Base
  - Content Migration
  - Preservation Standards Support

- Research Data
  - Integration with Research Systems
  - Deposit Tools
  - Sharing and Publishing

- Infrastructure
  - User Experience
  - Security
  - Performance & Scalability

© 2016 Ex Libris | Confidential & Proprietary
<table>
<thead>
<tr>
<th><strong>Infrastructure</strong></th>
<th><strong>DAM</strong></th>
<th><strong>Preservation</strong></th>
<th><strong>Research Data</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Support for Oracle 12c and Red Hat 7</td>
<td>Two-way Metadata Synchronization over OAI-PMH</td>
<td>Format Identification</td>
<td>CRUD Access Rights and Virtual Reading Room APIs</td>
</tr>
<tr>
<td>Upgrade to Java 8, BIRT 4.5, JQuery 2.1.4, Solr 5.2.1, Tomcat 7.0.65</td>
<td>Publish Digital Collections to Primo</td>
<td>Auto Correction Rules</td>
<td>Scheduled Sending of Reports</td>
</tr>
<tr>
<td>Enhanced Security</td>
<td>Mobile Support for Delivery</td>
<td>Index Technical Analyst Rules Events</td>
<td>Google Scholar Integration</td>
</tr>
<tr>
<td>Performance Improvements</td>
<td>Improved Collection Management</td>
<td>Preservation Plan Format ID Validation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CSV Deposit Enhancements</td>
<td>Droid 1.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Delivery Caching Support</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Rosetta – from a Point Solution to an Integrated Solution

A DAM and digital preservation solution interwoven into users workflows & institutional goals
Main Focus Areas Since Last IGeLU

1. Enhance Core Functionalities
2. Increase Synergy with other Ex Libris Products
3. Streamline the Implementation Process
The Power of the Best of Class Suite Approach

Market leading Discovery Solutions
Fully integrated with Rosetta – Digital collection browsing

Primo & Summon

Alma

Market Leading Library Management Solution
Fully integrated with Primo, Summon and Rosetta

Rosetta

Market Leading DAM and Preservation Solution
Integrated workflows with Alma and Primo
Two-way Metadata Synchronization

Manage Digital Objects Metadata in Alma using the familiar Alma UI

Seamlessly Preserve Objects’ Data and Metadata in Rosetta

Automatic Metadata Synchronization
Discover Digital Collections via Primo

Manage Digital Collections in Alma or Rosetta

Discover and Easily Navigate through Collections in Primo
Main Focus Areas Since Last IGeLU

1. Enhance Core Functionalities
2. Increase Synergy with other Ex Libris Products
3. Streamline the Implementation Process
Digital isn’t Easy

• Digital workflows are challenging to implement
• Lots of moving pieces
• Migration, ongoing ingest flows
• More IT and development involvement
• Connectors, storage, throughput, infrastructure are consistent pieces of the puzzle
• New file formats, delivery and access technologies
... and more
Preparing for the Unknown

- Multiple formats
- Multiple metadata standards
- Multiple metadata extractors
- Multiple format identifiers
- Multiple virus scanners

- Multiple publisher types
- Multiple viewers
- Multiple storage providers
- Multiple access rights policies
- Multiple report types

- Multiple collection types
- Multiple data types
- Multiple institution types
- Multiple ILSs
- Multiple discovery systems

- Multiple migration rules
- Multiple browsers
- Multiple users
- Multiple user types
- Multiple processes

- Multiple representations
- Multiple material flows
- Multiple delivery rules
- Multiple migration plans
- Multiple risks
Open, Extendable and Integrative

Primo

OAI-PMH/ SRU/SRW
Publisher Plugins
Converter Registry Plugins

Metadata Extraction Plugins
Virus Check Plugins
Custom Fixity Plugins

ILS Synchronization

Voyager
Alma
Aleph
Other ILS

PUBLISH
DELIVER
MANAGE
DEPOSIT
PRESERVE

Repository Task Plugins
Comparison Plugins
Migration Plugins
Risk Extractor Plugins

Producer Management Web Services
Field Validator Plugins
Deposit Web Services
Decomposer Plugins

Collection Management Web Services
BIRT reports
Data Management Web Services
Startups Check Plugins
CRUD Access Rights Web Services
Access Rights Plugins
IE Update Web Services
Persistent Identifier Plugins

Converter Registry Plugins
Viewer Pre-Processor Plugins

Delivery Web Services

NetApp
IBM
EMC
Oracle

NFS server
Streamline the Implementation Process

- Newly-formed dedicated implementation team
- Revised all implementation processes and deliverables
- New recorded training materials
- Developments supporting implementation needs
Rosetta Implementation Flow

**Preparation**
- Installation
  - Preparation Tasks
  - Sizing recommendations
  - Hardware provision
  - Site readiness
  - Installation
  - Rosetta Overview session

**Implementation Phase**
- Phase 1: Test Setup
  - Sample data provision
  - Data Analysis
  - Initial Configuration
    - Material Flows Analysis
    - Material Flow #1
    - Material Flow #2
  - Sample data ingest
  - Review Configuration
  - Basic Training

- Phase 2: Production Setup
  - Production Configuration
    - Copy configuration from Test
    - Phase 2 Configuration
  - Test ingest
  - Final Review
  - Start operational ingest
  - Integrations (if included)
  - Acceptance Testing
  - Advanced Training/Workshop

- Phase 3: Production Ingest
  - Backups
  - Health Checks
  - Review open issues
  - Switch to Support

**Documents**
- Rosetta Welcome Kit
- Rosetta Installation Kit
- Rosetta Training Kit
- Installation Request Forms
- Sizing Questionnaire
- Rosetta Implementation Site Readiness Checklist
- Rosetta Implementation Form – Phase 1
- Rosetta Configuration Form – Phase 2
- Rosetta Implementation Sign-Off
- Rosetta Health Checklist
- Rosetta Switch to Support Form
- Rosetta Implementation Survey

**Ex Libris Support during Project**
Rosetta Project Deliverables

**Preparation**
- **Installation**
  - Basecamp site
  - Production Sizing Recommendation
  - Rosetta Installations: test and production

**Implementation Phase**
- **Phase 1: Test Setup**
  - Kick off meeting
  - Basic Training
  - Follow-up calls
  - Initial Configuration for workflows
  - Sample records ingested
  - Documentation on configured Material Flows

- **Phase 2: Production Setup**
  - Production Configuration
  - Integration Setup
  - Production ingest started
  - Acceptance Testing
  - Advanced Training/Workshop

- **Phase 3: Production Ingest**
  - Operational Rosetta
  - Switch to Support session

**Closure**
- 2 Rosetta servers configured as requested: test and production
- Reports from readiness diagnostic tool
- Load balancer Setup
- Rosetta Implementation Site Readiness Checklist

- Sample records
- Rosetta Configuration Form – Phase 1
- Review of initial configuration

- Rosetta Configuration Form – Phase 2
- Review of production configuration
- Backup Setup
- Rosetta Implementation Sign-Off

- Rosetta Health Checklist
- Rosetta Switch to Support Form
- Rosetta Implementation Survey
Rosetta Training Flow

Initial Contact
- Overview
  - Watch (recorded)
    - Session 1: Preservation System Overview / Rosetta Data Model
  - Become familiar with Rosetta Terminology
    - Rosetta Glossary

Kick-off Meeting
- Essential Training
  - Watch the recorded Sessions and Ask Questions during Functional Calls
    - Session 2: Material Flows / Producers
    - Functional Call - Topic: "Ingest"
    - Session 3: Finding a SIP / Assessing & Arranging Content / The Web Editor
    - Session 4: Rosetta Task Chains / SIP Routing / Technical Analyst Workbench
    - Functional call - Topic: "SIP Processing"
    - Session 5: Set, Process, and Collection Management
    - Session 6: User Administration & Publishing / Primo Integration
    - Functional call - Topic: "Data Management"
    - Session 7: Preservation Planning
    - Session 8: Rosetta Administration
    - Session 9: Rosetta SP / HF Installation
  - Use Rosetta Documentation Guides
  - Hands-on Practice in the Training Institution

Operational Mode
- Workshop
  - Gather Questions for Workshop Agenda Preparation
  - Attend Workshop
  - Review Advanced Topics

Project Completed
- Practice
  - Use actively your Rosetta Installation
  - Extended Training
  - New Release Education
Main Focus Areas Since Last IGeLU

1. Enhance Core Functionalities
2. Increase Synergy with other Ex Libris Products
3. Streamline the Implementation Process
Rosetta 7th Advisory Group Meeting

Rosetta and the Emerging Trends and Challenges in Digital Preservation
June 7-8, NYU, New York
50 participants from 28 institutions

Stay tuned for the next RAG meeting!
Rosetta User Group Working Groups

- Ingest Tools
- Collection Management
- Technical Analyst Workbench
- Preservation Planning
- Security
- Access Rights
- Consistency Checks
- Format Library
- Preservation Planning
- Format Library
- Preservation Planning
- Format Library
Rosetta User Group Working Groups Enhancements

- Implemented - V 4.X: 24
- Implemented - V 5.0.X: 7
- Implemented - V 5.1: 8
- Planned - Future Versions: 12
- Pending: 17
- Obsolete: 16
- Deferred: 25

Total: 109
Implemented (including v5.1): 58 (53%)
Roadmap Overview
# Improved User Experience

New UI introduced throughout the system
SIP Deposit Prioritization

Enable fast deposit of content with higher priority

Deposit Prioritization Settings:
- Validation Stages - Pre-approval Automatic Stage:
  - Validation Stack Routine: Validation Stack - Full
  - Approval Stage - Manual Approval/Assessment Stage:
    - Approval: Assessor and Arranger
    - Allow Split/Merge: Yes
- Enrichment - Post Approval Automatic Stage:
  - Enrichment Routine: Enrichment - Minimal

Done
Improved deposit tools providing better control on the deposit process

- Improved submission job history management
- Improved deposit error handling and logging
- Edit submissions via the web deposit UI
Validation Stack Enhancements

Technical Analyst Workbench

• Improved design
• Technical Analyst rules export to enable sharing rules within institutions

Virus Check Plug-in

• Multiple exit codes enabling error differentiation
• Rules for automatic handling of virus check errors

Fixity Checks

• OOB support for SHA-256

© 2016 Ex Libris | Confidential & Proprietary
Duplicates Management

Define criteria for duplicates

Identify duplicates on deposit

Remove duplicates from the system
Enhanced Search

- Support NOT in SOLR queries
- Improved smart-search for all controlled-vocabulary fields
- Improved indexing and search performance
Easy editing of text and XML files using a graphical XML editor:

- Structmaps
- Source metadata
- Configuration files
- XSL files

Example XSL transformation:
```xml
<xsl:stylesheet xmlns:mets="http://www.loc.gov/METS/"
    xmlns:dnx="http://www.exlibrisgroup.com/dps/dnx"
    xmlns:dc="http://purl.org/dc/elements/1.1/"
    xmlns:dcterms="http://purl.org/dc/terms/"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xmlns:xs1="http://www.w3.org/1999/XSL/Transform"
    xmlns:mods="http://www.loc.gov/mods/v3"
    version="1.0">
  <xsl:output method="xml" indent="yes" omit-xml-declaration="yes"/>
  <xsl:template match="/">
    <dc:record xmlns:dc="http://purl.org/dc/elements/1.1/"
                xmlns:dcterms="http://purl.org/dc/terms/"
                xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
                xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    <xsl:if test="/dc:dc/dc:dc">
      <xsl:copy-of select="/dc:dc/dc:*/"/>
    </xsl:if>
  </dc:record>
</xsl:stylesheet>
```
Enhanced Source MD Support

- Edit Source MD using a graphical XML editor
- Index Source MD elements
- Enable searching through source MD elements
Improved Collection Management

Collection Management UI

- Collection ordering
- New collection history tab

Automatic Collection Management

- Create and update collections thumbnails via APIs
- Collection ordering API
- New task to assign collection by METS
- Automatically unassign IEs from sub-collections

Collection Discovery

- Discover collections in Primo based on the given order
New plug-in to automatically determine the quality of migration plans by comparing between files pre- and post-migration.
Preservation Enhancements

Consistency Checks
- New report for retrieving objects’ full storage path
- New plug-in for identifying missing files

Simplified Preservation Planning
- Enable creating a preservation plan from more than one file format

Custom Events Support
- Enable depositing of provenance events into Rosetta
- Custom events will be searchable, publishable, and exportable

Tracking Changes to Access Rights Policies
- New AR policy audit trail page that will track policy changes
- Provide APIs for tracking changes
Enable publishing of derivative copies - and not only preservation representations - to external discovery systems using OAI-PMH
Enhanced Viewers

New Viewers

- New JWPlayer with support for HTML5
- New FlexPaper with support for HTML5
- New REP viewer supporting native HTML5 players
- Additional viewers are currently under review

Viewers Customization

- Enable setting any viewer parameters from within Rosetta
- Allow external viewers to be called outside of Rosetta’s delivery page context
New APIs

### IE Update
- Add derivative copies
- Update representation's non-physical StructMap
- Get file physical location
- CRUD access rights on access copies

### Deposit
- Get SIP Status by External ID
- SIP Status for List of SIPs

### Collection Management
- Create and update collection thumbnail
- Collection ordering

### Delivery
- Access rights check service
Integration with Alma

- Alma Integration Plug-in - handles Remote D inventory CRUD operations
- Submission application to create an Alma digitization request and generate a SIP accordingly
- Preservation for Alma Digital content (metadata and files managed in Alma)
- Access Rights check for remote digital repository
SWORD and BagIt Support

**SWORD**
- Support depositing data from SWORD repositories into Rosetta

**BagIt**
- Support depositing BagIt "bags" into Rosetta
- Support exporting BagIt "bags" from Rosetta
Performance Improvements

- Deposit and Ingest
- Search and Indexing
- Publishing
- Data Processing
Support Single Sign-on (SSO) over SAML (Security Assertion Markup Language)

SAML Identity Provider (Customer Hosted)

User

Rosetta - SAML Service Provider
Documentation and Training

- New training videos

- Improved plug-ins and APIs documentation

- Provide documentation for all Rosetta’s events

- Provide documentation for email notifications mapping

- Load Balancer documentation

- Revised Rosetta-Primo Integration Guide
Version 5.1 Highlights – September 2016

- Enhanced Viewers Customization
- New Viewers
- Enhanced Search
- Collection Ordering
- SIP Deposit Prioritization
- Improved Deposit Management
- Derivative Copies Publishing and Creating
- New XML Editor
- Improved Collection Management

- DAM
- Research Data
- Preservation

- Create a Preservation Plan for more than One Source Format
- Share Technical Analyst Rules between Institutions
- File Comparison Plug-in
- Consistency Checks
- Tracking Changes to Access Rights Policies
- Validation Stack Enhancements

| Infrastructure | Performance Improvements | Enhanced Clean-up Processes | Enhanced Indexing | New UI | New APIs |
Questions?
THANK YOU!
Adi.Alter@exlibrisgroup.com