Rosetta Product Update

IGeLU 2017

Adi Alter | Director of Product Management, Digital Solutions
Status Update
Main Focus Areas Since Last IGeLU

1. Enhance Core Functionalities
2. Increase Synergy with Other Ex Libris Products
3. Streamline the Implementation Process
From a Point Solution to an Integrated Solution

- Digital Preservation 2007
- Digital Preservation and DAM 2012
- A DAM and Digital Preservation Solution Interwoven into Users Workflows & Institutional Goals 2016
Enabling Rosetta to Serve as the Institutional Digital Repository Backbone

Rosetta as the Institutional Digital Repository Backbone
Enhance Core Functionalities

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

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

Infrastructure
- User Experience
- Security
- Performance & Scalability
Versions 5.1 Highlights – September 2016

**Digital Asset Management**
- SIP Deposit Prioritization
- Improved Deposit Management
- Enhanced Search – ‘Not’ Operator Support and Smart Search
- Collection Ordering and History
- Improved Collection Management
- New Graphical XML Editor
- Derivative Copies Publishing and Creation
- New HTML5 Viewers
- Enhanced Viewers Customization

**Preservation**
- Create a Preservation Plan for more than One Source Format
- Share Technical Analyst Rules Between Institutions
- File Comparison Plug-in
- Consistency Checks
- Validation Stack Enhancements - SHA-256 Fixity, Virus Check Exit Codes and Automatic Rules

**Infrastructure**
- Update Software Stack - Tomcat, PDS, Aurigma Uploader
- New User Interface
- Performance Improvements
- Enhanced Indexing
- New APIs

© 2017 Ex Libris | Confidential & Proprietary
Versions 5.2 Highlights – February 2017

Digital Asset Management
- Alma Digital Inventory’ Repository Tasks
- Alma Digitization Submission Application
- Access Rights Check for Remote Repositories
- OAI-PMH Harvester Test UI
- Improved TA Workbench and TA Error Visibility
- Submission Job Reports
- Search and SRU Improvements
- Delete Derivative Copy Representation Task
- Full Text Viewer Pre-Processor (VPP)

Preservation
- Deposit Content using BagIt ‘Bags’
- Retaining Timestamps for Digital Content Uploaded to Rosetta
- Storage Migration
- Indexing Support for Long Fields
- Access Rights Policies Audit Trail

Infrastructure
- SAML / Local Authentication
- Improved User Experience
- Performance Improvements
- Enhanced Clean-Up Jobs
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
Workflow I: Two-Way Metadata Synchronization Using OAI-PMH

1. Deposit
2. Store
3. Publish (Rosetta) / Harvest (Alma) (OAI-PMH)
4. Create Digital Representation / Bib Records
5. Manage Metadata
6. Publish (Alma) / Harvest (Rosetta) (OAI-PMH)
7. Update Metadata
8. Publish (OAI-PMH)
9. View
The Metadata in Alma and Rosetta

MARC Record in Alma

DC Record in Rosetta
Workflow II: Digitization Requests’ Processing

1. Create Digitization Request
2. Deposit
3. Store
4. Create Remote REP (API)
5. Close Digitization Request (API)
Digitization Request Fulfillment

 Pakistani women : pakistani feminists, women's rights in pakistan, aafia siddiqui, women in.

- **Language**: English
- **ISBN**: 1230580697 and others
- **Record number**: (OCoLC)867836167

**Subject**: Pakistan and women's rights

**Digitization Options**
- Physical (1)
- Electronic
- **Digital (1)**

**Collections (1)**
Workflow III: Digitization – Alma Submits to Rosetta (Roadmap)

1. Create Digitization Request

2. Deposit

3. Create Digital Representation / Bib Records

4. Deposit (API)

5. Store

6. Close Digitization Request (API)
Workflow IV: Rosetta Preserves Alma Digital (Roadmap)

1. Deposit
2. Create Digital Representation / Bib Records
3. Store
4. Manage Metadata
5. Publish (Alma) / Harvest (Rosetta) (OAI-PMH)
6. Store
7. Publish (OAI-PMH)
8. View

© 2017 Ex Libris | Confidential & Proprietary
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

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

- Primo

- Metadata Extraction Plugins
- Virus Check Plugins
- Custom Fixity Plugins
- ILS Synchronization

- Voyager
- Alma
- Aleph
- Other ILS

- NetApp
- IBM
- EMC
- NFS server

- Storage Abstraction Layer

- Repository Task Plugins
- Comparison Plugins
- Risk Extractor Plugins
- Migration Plugins

- PUBLISH
- DELIVER
- INGEST
- MANAGE
- PRESERVE
- DEPOSIT

- 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

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

- Delivery Web Services
- Viewer Pre-Processor Plugins
Out-of-the-Box Tools & Configurations

- Deposit flows
- Task chains
- Delivery rules
- Plug-ins
- Format library
- Access rights configuration
- User roles
- BIRT reports
Constantly Enhance Rosetta’s Out-of-the-Box Offering
Streamline the Implementation Process

- Dedicated cross-regions Rosetta 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

**Closure**
- 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

### Customer
- 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

---

© 2017 Ex Libris | Confidential & Proprietary
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 for the Upcoming Year

1. Enhance Interoperability and OOB Integrations and Tools

2. Improve User Experience and System Performance

3. Enrich Core Functionalities Through Streamlined Collaboration with the RUG
Continue Investing in an Open Architecture
While Enhancing Rosetta’s Out-of-the-Box Tools’ Offering
Improve User Experience and System Performance

- Continue the New UI Design Uplift
- Simplify Workflows
- Support User Preferences

PROVIDING A SEAMLESS USER EXPERIENCE

- A New Dedicated Infrastructure Team
- Enhanced Performance Lab
- New System Monitoring Tools

ROBUST, EFFICIENT & SCALABLE DIGITAL PRESERVATION PLATFORM
Enrich Core Functionalities Through Streamlined Collaboration with the Rosetta User Group

- Delivery & Integrations
- Digital Preservation
- Format Library
- System Operations
Rosetta Version 5.3 – August 2017
Version 5.3 Highlights – August 2017

Digital Asset Management
- Native IIIF Support
- SIP Metadata Search
- Duplicates Report
- Collection Display Management
- AIP Update API Enhancements
- OAI-PMH Harvesting -Matching by Other IDs

Preservation
- Store TA Rule ID as Part of the TA Rule Event
- Store Identification Method in File Format Section in DNX
- Keep Original File Format Identification Event in DNX
- Event History Report

Infrastructure
- Copy Configuration
- Improved User Experience
- Drag-and-Drop Rows in Tables
- Performance Improvements
- Support for Multiple User Groups
- Start Rosetta in Safe Mode
SIP Metadata Search

• SIP metadata is indexed early in the ingest process enabling searching through ingested content in different stages (TA Workbench, 3As)

• This metadata will be searchable during the ingest process via
  • A new search screen
  • SRU
Duplicates Report

• A new job to identify duplicate objects by:
  • Checksum value
  • Metadata identifier:
    • dc:identifier
    • dcterms:identifier
    • object identifier

• Both on the consortium and institution levels
Event History Report

• Enables users to generate a report by any event:
  • Takes an event ID and date range
  • Returns a list of PIDs, event date, and event key values

• Examples:
  • Indexing events (to track failed indexing)
  • CRUD reporting for set and process management
figshare Integration

- Enhance the OAI-PMH harvester to match records of external origin based on a dc:identifier or an IE ObjectIdentifier

- Repository Task plug-in to notify figshare upon AIP creation
The **International Image Interoperability Framework** (IIIF) provides a standardized method of describing and delivering images over the web by defining the following APIs:

- **Image API** – provides the image with characteristics such as region, size, rotation, quality and format
- **Presentation API** - describes the structure and layout of a digitized object
- **Search API** - allows for "searching annotation content within a single IIIF resource"

These APIs can be used to develop image servers and web clients that provide a world-class user experience in viewing, comparing, manipulating and annotating images.
IIIF Support - What Will Rosetta Do For You?

• **IIIF Manifest**
  • Describes the file structure and contains references to Image API endpoints.

• **Out-of-the-box IIIF Server - Cantaloupe**
  • An open-source image server with operations such as cropping, scaling, and rotation, it can support deep-zooming image viewers, as well as on-the-fly thumbnail generation.

• **Out-of-the-box IIIF Viewer – Universal Viewer**
  • An open-source viewer with a wide range of formats supported providing advanced features such as zooming, navigation and more.

• **Out-of-the-box File Viewer – OpenSeadragon**
  • An open-source, web-based viewer for high-resolution zoomable images.

It will be possible for institutions to use all components provided by Ex Libris or to take some of them and use their own server or client.

Support for **IIIF Open Annotation Extensions** - 2018
Version 5.3 Highlights – August 2017

Digital Asset Management
- Native IIIF Support
- SIP Metadata Search
- Duplicates Report
- Collection Display Management
- AIP Update API Enhancements
- OAI-PMH Harvesting - Matching by Other IDs

Preservation
- Store TA Rule ID as Part of the TA Rule Event
- Store Identification Method in File Format Section in DNX
- Keep Original File Format Identification Event in DNX
- Event History Report

Infrastructure
- Copy Configuration
- Improved User Experience
- Drag-and-Drop Rows in Tables
- Performance Improvements
- Support for Multiple User Groups
- Start Rosetta in Safe Mode
Rosetta Roadmap Overview
Rosetta Roadmap Schedule

Version 5.4  December 2017

Version 6.0  June 2018

Version 6.1  December 2018

Version 6.2  June 2019
Version 5.4 Highlights – December 2017

**Digital Asset Management**
- Extend SRU Capabilities
- Renewed IE / METS Viewer
- Download all IE Files
- IA Book Reader as a IIIF Viewer
- Enable Sending Emails for Jobs Statuses
- Enhance the UI with Process Automation Logs

**Preservation**
- TA Workbench – Bulk Operations
- TA Workbench – Manual Format and Extension Assignment
- Out-of-the-Box Format Identifiers
- Out-of-the-Box Tech MD Extractors
- Out-of-the-Box Migration Tools

**Infrastructure**
- Update Software Stack - Solr 6.0 (Solr as a Server), BIRT 4.7, Tomcat 8.5
- New BIRT Reports UI
- Performance Improvements
TA Workbench Improvements

• Enable selecting multiple files within a SIP and applying a single action on them
  • e.g. assign file extension, assign format, ignore error, etc.

• Enables ad-hoc rules for the current SIP

• Manual format and extension assignment
Out-of-the-Box Preservation Tools

• Enhance Rosetta’s out-of-the-box preservation tools with tools *such as*:

  • **Metadata Extractors (pending FLWG):**
    • *ExifTool* – for image, audio, video, and PDF content
    • *MediaInfo* – for media content

  • **Validation Tools**
    • *veraPDF* – for PDF files

  • **Migration Tools**
    • *FFMPEG* - for multimedia content
Improved IE / METS Viewer

Configurable, responsive, new UI similar look and feel as the recently added REP viewer
Extend SRU Capabilities

• Add Authentication to SRU/SRW Requests

• Add additional query operators such as ‘<’, ‘>’, ‘!=’, etc.
• Add query operators between fields – such as ‘or’ and ‘not’
• Add ability to search by partial date (e.g. year only)

• Enable configuration of the SRU response with an xsl file
## Version 5.4 Highlights – December 2017

### Digital Asset Management
- Extend SRU Capabilities
- Renewed IE / METS Viewer
- Download all IE Files
- IA Book Reader as a IIIF Viewer
- Enable Sending Emails for Jobs Statuses
- Enhance the UI with Process Automation Logs

### Preservation
- TA Workbench – Bulk Operations
- TA Workbench – Manual Format and Extension Assignment
- Out-of-the-Box Format Identifiers
- Out-of-the-Box Tech MD Extractors
- Out-of-the-Box Migration Tools

### Infrastructure
- Update Software Stack - Solr 6.0 (Solr as a Server), BIRT 4.7, Tomcat 8.5
- New BIRT Reports UI
- Performance Improvements
## Version 6.X Highlights (Initial List)

### Digital Asset Management
- Export to FTP
- Delete OAI Sets
- Extended IIIF Support
- Improved Collection Management
- Access Rights on Files / Structmaps
- New APIs
- New Reports

### Preservation
- Support Multiple MD Extractors
- Support Multiple Format Identification Tools
- Support Custom Events
- Add Provenance Information about Original Checksums
- Improved Custom Fixity Support
- TA Workbench Improvements
- BagIt Support – Export
- Out-of-the-Box Tech MD Extractors
- Out-of-the-Box Migration Tools
- Improved Preservation Flow

### Infrastructure
- Support for Oracle 12.2
- User Preferences Support
- Favorite Menu Items
- Performance Improvements
THANK YOU!
Adi.Alter@exlibrisgroup.com