**Rosetta's Ingest Methods –
What's the Right Path to Take?**

Zvi Fass
Rosetta Development Team Leader

ExLibris
a ProQuest Company

Introductions

- Zvi Fass - Rosetta Development Team Leader

    - In Ex Libris since 2006
    - Part of the Rosetta Development team since day one
    - Development Team Leader for the past two years

ExLibris

Objectives and Target Audience

- Brief Session Description:
  - We will review different Rosetta ingest methods – such as Web deposit, CSV, METS, OAI-PMH, BagIt and Deposit APIs - along with the advantages & limitations of each method.

- Session Objective(s)
  - **What's the right path to take**? By the end of the session you will be able to decide what's the best path to take with your specific data and needs.

- Session Target Audience:
  - All Rosetta users.

3   **ExLibris**

Agenda

1 Introduction

2 Manual Ingest

3 Automated Ingest

4 Bulk Ingest

5 Next Steps, Support Resources and Survey

4   **ExLibris**
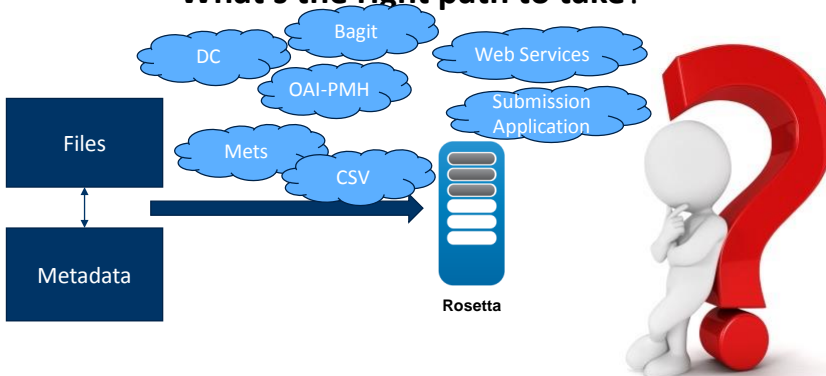
Introduction

5

ExLibris

Introduction

Depositing into Rosetta
What's the right path to take?

Bagit
DC
Web Services
OAI-PMH
Submission Application

Files
Mets
CSV

Metadata

Rosetta

6

ExLibris

3

Introduction

Basic terminology
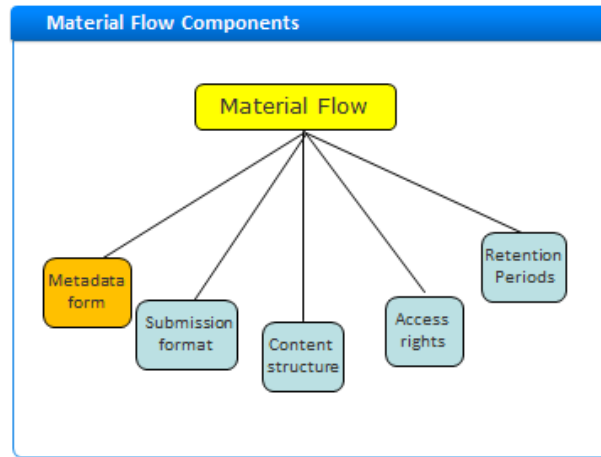
Material Flow Components

ExLibris

Agenda

1 ~~Introduction~~

2 Manual Ingest

3 Automated Ingest

4 Bulk Ingest

5 Next Steps, Support Resources and Survey

ExLibris

Manual Ingest

9

Manual Ingest

Manual Material Flow

Advantages	Limitations

• User friendly – wizard
• Customizable Metadata Form
• Customizable Submission Format
   • Bulk uploader
   • Http
   • Detailed
	• Non bulk
	• Only supports one Rep per IE
	• Content Structure - IE per file or single IE
	• Only PC upload
	• No DNX Metadata

10

Agenda

1 ~~Introduction~~

2 ~~Manual Ingest~~

3 Automated Ingest

4 Bulk Ingest

5 Next Steps, Support Resources and Survey

ExLibris

Automated Ingest

ExLibris

Automated Ingest

Automated Material Flow

- Non web ingest (can also be done via web)

- Supports multiple Representations in multiple IEs.

- Customizable Submission Format
  - NFS
  - FTP
  - SFTP

- Customizable Content Structures
  - Dublin core Converter
  - Mets Converter
  - CSV Loader Converter
  - Bagit Converter

13

ExLibris

Automated Ingest

Dublin Core Converter

Advantages	Limitations
• Automated – Bulk Ingest • Simple - Need only DC or no metadata • Support of multiple IEs • Relative/Absolute/Http Paths	• Supports only DC metadata • No DNX metadata • No Source metadata • No support of multiple representations

14

ExLibris

Automated Ingest

Mets Converter

Advantages	Limitations
• Automated – Bulk Ingest • Multiple IEs support • Multiple Representations support • Relative/Absolute/Http Paths • Metadata update – full flexibility • Collection support (assignment only)	• Complex • Need a Submission Application to prepare the Mets for Ingest

ExLibris

Automated Ingest

CSV loader Converter

Advantages	Limitations
• Automated – Bulk Ingest • Multiple IEs support • Multiple Representations support • Relative/Absolute/Http Paths • Metadata update – full flexibility • Full Collection support • Simple to create	• Need a tool to create the CSV Files • No support for Logical Struct Map

ExLibris

Automated Ingest

Bagit Converter

Advantages	Limitations
• Automated – Bulk Ingest • **BagIt** format support • IE DC and DNX metadata • Rep DC and DNX metadata	• No metadata support for files • No Logical Struct Map support

ExLibris

Agenda

1 Introduction

2 Manual Ingest

3 Automated Ingest

4 Bulk Ingest

5 Next Steps, Support Resources and Survey

ExLibris

Bulk Ingest

19

Bulk Ingest – Submission Job

- Pre-prepared load packages (according to the Material Flow)
- Automatic scheduling
- Locking mechanism

20

Bulk Ingest – OAI-PMH Harvester

Rosetta

Get OAI-PMH Records (HTTP)

External
Repository

Transform / Match

File System

Generate Submission Job /
Update Metadata Job* Folders

ExLibris
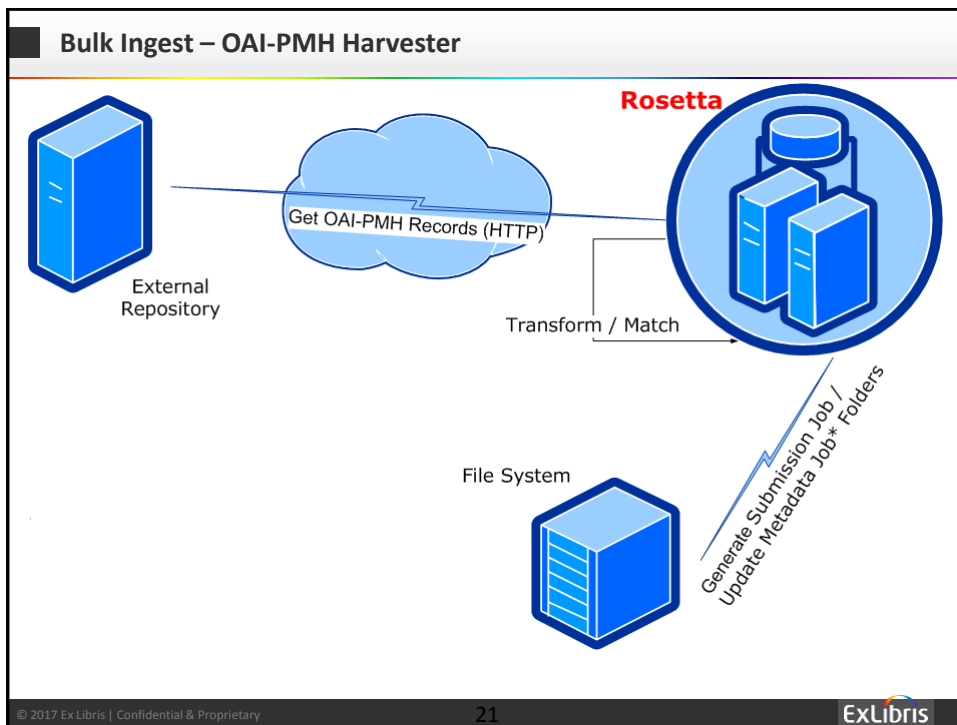
Bulk Ingest – OAI-PMH Harvester

The end-to-end process of harvesting records is comprised of
four stages:

1. Harvesting
2. Matching
3. Transformation
4. Based on the matching results generates a:
   - Submission Job folder for new records
   - Metadata Update folder for existing records.

Scheduled Submission and/or Metadata Update Jobs run
independently of the Harvesting Job.

ExLibris

Bulk Ingest – OAI-PMH Harvester

Bulk Ingest - Deposit Web Services

Deposit Web Services (wsdl)

Allow institutions to develop applications over Rosetta

- Customers would like to use their own deposit frontend, which includes creation of IEs, depositing and retrieving deposit Information (for example Indigo)
- Bulk Deposits – creation of automated tools to deposit massive numbers of IEs
- Migrate existing metadata into Rosetta

Bulk Ingest - Deposit Web Services

Submission Application

- SDK – development tools
- IE Creation – Mets deposit (or any other structure)
- Deposit – using the Deposit web services
- Monitoring - using the SIP web services

25
ExLibris

Bulk Ingest - Submission Application

Six Step Submission Application

1. Create a SIP directory
2. Create a Rosetta Compatible Mets
3. Synchronize Deposit Directory with Submission Format
4. Authentication with PDS/Header authentication *New*
5. Submit Deposit
6. Retrieve the SIP status

The source code of the example can be found under:
dps-sdk-<version>\src\com\exlibris\dps\sdk\examples\FullFlowExample. java

26
ExLibris

Recap

	Web	Bulk	DC Update	DNX IE/REP	DNX File	Source MD	Collections	Struct Map
Manual	✓	✗	✓	✗	✗	✓	✗	✗
Dublin Core Converter	✗	✓	✓	✗	✗	✗	✗	✗
Mets Converter	✗	✓	✓	✓	✓	✓	✓ (only assign)	✓
CSV loader Converter	✓	✓	✓	✓	✓	✓	✓	✗
Bagit Converter	✗	✓	✓	✓	✗	✗	✗	✗

Recap

**Depositing into Rosetta
What's the right path to take?**

What's Your Business?

Migration

Ingest Flows

End user

4/20/2017

Next Steps and Support Resources

© 2017 Ex Libris | Confidential & Proprietary

ExLibris

Next Steps and Support Resources

- Include documentation links relative to the topic:
  - SDK
  - Deposit Web Services
  - OAI-PMH Harvesting

- Additional support resources within the ExLibris Ecosystem:
  - Idea Exchange
  - Developer Network

- Technical Seminar Presentations
  (located in the Cross-Product section of the CKC)

© 2017 Ex Libris | Confidential & Proprietary

ExLibris

Q & A

Any Questions?

31

ExLibris

Session Survey Evaluation

Please use the following link
https://www.surveymonkey.com/r/techsem2017
to provide feedback on your sessions.

32

ExLibris

THANK YOU