Primo & Summon: Unified Content Approach
AND HOW IT WILL BENEFIT YOUR USERS

Christine Stohn | Product Manager
Eddie Neuwirth | Product Manager
Key Topics

• Data-based services across Summon and Primo

• Shared content and data model

• Unified content ingestion process

• Discussion - Q/A
Primo Central and Summon content – Fun Facts

~13 TERABYTES OF DATA

- More than 500 provider agreements
- Combined more than 4 Billion documents

IF PRINTED COULD GO AROUND THE WORLD 19 TIMES

COULD MAKE 780 STACKS OF PAPER AS HIGH AS THE EIFFEL TOWER
Data-based services
Data for discovery

- More content, shared processes
- Better glanceability
- Content neutrality
- More data-based services
- More data
- More data connections
- Collaborative research
Data for exploration services

Citation Trail

Topic Exploration

Article Recommender (bX)

DB/ Resource Recommender
Exploration through Data-based Services: Future considerations

- Connections and linked data (e.g. via ORCID, ISNI)
- Open Access discovery
- Visualization
- Data correlations
No change to the method of managing local collections in Primo and Summon

>90 Content types  >4,000 Databases  >2.7 B records
Streamlined process for adding new content

Increased velocity for getting new and backlog content indexed

Provides opportunities to further enrich metadata and correct issues with metadata
What it all means for each service

Content    Metadata

Content    Metadata

ExLibris Primo

Summon

New content
First Let’s Call the Lawyers!

More than 500 Agreements with Publishers to Review

Mid-August 2016 Discovery Dashboard

**2016 Primo to Summon Rights**
- Agreements signed:
  - Q2: 30
  - Q3: 20
  - Q4: 13

**Status: On Track**
- 2016 Goal: 63 agreements
- Signed YTD: 35 agreements

**2016 Summon to Primo Rights**
- Agreements signed:
  - Q2: 50
  - Q3: 40
  - Q4: 17

**Status: On Track**
- 2016 Goal: 107 agreements
- Signed YTD: 61 agreements
First collection completed in May!

• American Institute of Aeronautics and Astronautics (AIAA) now in Primo
  - More than 300 books, articles from 16 journals, 5 proceedings titles and also standards - no item level count available for now

• Increased velocity after that date through approx. June 2017
Shared Content So Far (As of September 1)

> > 150 Collections shared from Summon to Primo

### Shared Collection Examples

<table>
<thead>
<tr>
<th>Collection Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vandenhoeck &amp; Ruprecht</strong></td>
<td>Journals and ebooks</td>
</tr>
<tr>
<td><strong>Knovel Corporation</strong></td>
<td>ebooks</td>
</tr>
<tr>
<td><strong>American Medical Association</strong></td>
<td>Journals</td>
</tr>
<tr>
<td><strong>ACM Digital Library</strong></td>
<td>Conference proceedings, journals</td>
</tr>
<tr>
<td><strong>Al Manhal FZ</strong></td>
<td>Journals, books, dissertations</td>
</tr>
<tr>
<td><strong>UKSG</strong></td>
<td>Journals</td>
</tr>
<tr>
<td><strong>Brill</strong></td>
<td>ebooks and journals CRC Press</td>
</tr>
<tr>
<td><strong>Australian National Data Service</strong></td>
<td>Research from Australian research organizations</td>
</tr>
<tr>
<td><strong>Chinese Academy of Social Sciences (CASS)</strong></td>
<td>Chinese Academy of Social Sciences research</td>
</tr>
<tr>
<td><strong>Defense Technical Information Center</strong></td>
<td>DTIC Technical Reports</td>
</tr>
<tr>
<td><strong>Oxford University Press</strong></td>
<td>Scholarly editions content</td>
</tr>
<tr>
<td><strong>Cambridge University Press</strong></td>
<td>Cambridge Histories</td>
</tr>
<tr>
<td><strong>American Institute of Physics</strong></td>
<td>Journals, Open Access</td>
</tr>
<tr>
<td><strong>Optical Society of America</strong></td>
<td>Journals, Reports</td>
</tr>
<tr>
<td><strong>Igaku Shoin</strong></td>
<td>Medical Finder Journals</td>
</tr>
<tr>
<td><strong>Key Note</strong></td>
<td>Market Research reports</td>
</tr>
<tr>
<td><strong>Equinox Publishing</strong></td>
<td>Journals</td>
</tr>
<tr>
<td><strong>Tanner Ritchie Publishing</strong></td>
<td>Medieval and Early Modern Sources Online (MEMSO)</td>
</tr>
<tr>
<td><strong>Nurimedia</strong></td>
<td>ebooks</td>
</tr>
<tr>
<td><strong>American Museum of Natural History</strong></td>
<td>Research Library, Reports</td>
</tr>
</tbody>
</table>

© 2016 Ex Libris | Confidential & Proprietary
Historical View & Growth Rate

Approx. 9% growth so far in 2016

Record Count

- 2010: 500,000,000
- 2011: 1,000,000,000
- 2012: 1,500,000,000
- 2013: 2,000,000,000
- 2014: 2,500,000,000
- 2015: 3,000,000,000

Index growth

- 2010: 32%
- 2011: 29%
- 2012: 25%
- 2013: 16%
- 2014: 18%
- 2015: 18%

Approx. 9% growth so far in 2016
Unique Content to Share - Summon

• Summon has over 800 digital collections and institutional repositories available to share
  • Much of this content is library contributed and unique special collections
  • Most of it is open access
Unique Content to Share – Primo

• Primo Open Access
  • 40+ collections
  • OA flag for hybrid

Have you heard of recent requirements by the National Institutes of Health (NIH), as well as pending requirements for National Science Foundation (NSF) grant recipients to have a plan in place for data management?

Do you know your University Libraries has a repository service available to you that addresses these needs?

ThinkTech is an online publishing and archival service for the Texas Tech University community that scholars may use to distribute and archive their scholarly works with no limitations on content or format. ThinkTech also provides access to materials that traditional publications cannot. Users may access:

- Gray literature
- Video
- Audio
- Images
- Data sets
- Charts/graphs
Support for Open Access Discovery

New OA flag in the Summon schema to support Open Access initiatives for Summon and Primo
Unified Content Ingestion Process

- Summon ingest process, normalization, metadata clean-up and enrichment

Shared electronic collections

Loading of existing content and data enrichment to Primo Central / Summon Index
The Content Ingestion Process for Commercial Content

Identify New Content to Add into Discovery
Identifying New Commercial Content

• Product Management, Sales, and our Global Content Alliance work together to identify new content to add into Summon & Primo

• New content requests from customers:
  • NERs voting
  • Idea Exchange

• Publishers and content providers may also request to have their content added
The Content Ingestion Process for Commercial Content

1. Identify New Content to Add into Summon
2. Pre-Agreement Content Sample Analysis
3. Engage with Publisher/Provider
Pre-Agreement Content Sample Analysis

• The sample analysis is used to help determine the quality and extent of the metadata and the metadata schema

• We also try to determine things such as linking methods, how rights are assigned to the content, and what databases we may need in our knowledgebase (if they don’t already exist)

• We often indexes content at the article-level, or chapter-level as that is usually the level of granularity that the content is supplied at
What Metadata Do We Look For During Sample Analysis?

**Title Metadata**
- Article titles, Book titles, Publication titles, Subtitles, etc.

**Identifier Metadata**
- Unique IDs for specific articles, chapters, etc.
- Publication-level unique identifiers such as ISSN or ISBN
- Additional identifiers such as OCLC Number, LCCN, DOI, etc.

**Publication Information Metadata**
- Publisher, Author(s), Corporate Authors, Volume Numbers, Issue Numbers, Start Page, Publication Date, Publication Series, ORCID etc.

**Additional Metadata**
- Subject Headings, Keywords, Language, Classification
The Content Ingestion Process for Commercial Content

1. Identify New Content to Add into Summon
2. Engage with Publisher/Provider
3. Pre-Agreement Content Sample Analysis
4. Formalize and Sign Data Sharing Agreement
5. Data is Delivered in Full from Publisher/Provider
The Content Ingestion Process for Commercial Content

1. Identify New Content to Add into Summon
2. Engage with Publisher/Provider
3. Pre-Agreement Content Sample Analysis
4. Formalize and Sign Data Sharing Agreement
5. Data is Delivered in Full from Publisher/Provider
6. Data Normalization, Mapping, and Enrichment
Data Normalization, Mapping, and Enrichment Work

Data Normalization

• Very basic high-level clean-up of the data to standardize it
• Examples include:
  • Remove leading/trailing white spaces in Title and Subtitle fields
  • Clean-up diacritics and other encoding issues

Mapping

• Map the metadata fields in the records to the Summon schema
• This allows the metadata to appear in the UI and/or be made searchable within Summon (and Primo)

Enrichment

• Enriching the content by adding additional metadata when applicable
• Examples:
  • Scholarly/peer-reviewed flags from Ulrich’s
  • Citation counts from Scopus
  • Book cover images from Syndetics
The Content Ingestion Process for Commercial Content

1. Identify New Content to Add into Summon
2. Engage with Publisher/Provider
3. Pre-Agreement Content Sample Analysis
4. Formalize and Sign Data Sharing Agreement
5. Data is Delivered in Full from Publisher/Provider
6. Data Normalization, Mapping, and Enrichment
7. Indexing
8. Match-Merge Super Record Creation
9. Preferred Record Enrichment

Ex Libris Primo
Summon
Record Deduplication: Merged “Super Records”

- Summon displays a Single Record
- Primo displays an enriched record and options to view all versions

Cancer: The RAS Renaissance
by Ledford, Heidi

*Nature*, ISSN 0028-0836, 04/2015, Volume 520, Issue 7547, pp. 278-280
The Content Ingestion Process for Commercial Content

1. Identify New Content to Add into Summon
2. Engage with Publisher/Provider
3. Pre-Agreement Content Sample Analysis
4. Formalize and Sign Data Sharing Agreement
5. Data is Delivered in Full from Publisher/Provider
6. Data Normalization, Mapping, and Enrichment
7. Indexing
8. Post-Ingestion Maintenance

- Preferred Record Enrichment
- Match-Merge Super Record creation
Post-Ingestion Maintenance

Currency

• Currency is the process of the publisher/provider sending to Summon new/updated metadata records, or record deletions for content that need to be removed

• Frequency of providing updates is often at the discretion of the publisher/provider

Update Mappings

• For library catalog and institutional repository content, update mappings upon request

• For commercial content, will update mappings as needed or if issues are reported

Metadata Issues

• Address reported issues of metadata quality

• Most issues involve incorrect metadata, or slight variations in the metadata that may impact OpenURL linking or the record deduplication process

• We do not make metadata corrections to library catalog or institutional repository records
Great Services Fueled by Shared Content and Data Expertise

- Most comprehensive central index and always growing...
- Unique collections
- Commitment to content neutrality
Discussion / Q&A

THANK YOU
eddie.neuwirth@exlibrisgroup.com
christine.stohn@exlibrisgroup.com