One Object – Three Systems

DoD with ALEPH, DigiTool and Primo

Dirk Willinghöfer
The National Library & Ex Libris

- **ALEPH**
  - Since 1999
  - Participant in Austrian Library Consortia
- **DigiTool**
  - Since 2004/5
The National Library & Ex Libris

- Primo („Quicksearch“)
  - Since 2011 (stp May 2011)
  - Consortial installation hosted by OBV
Digitization On Demand

• Copy Requests
  – Since 2009 all copy requests by users are handled digitally
  – Department decides whether to scan the full item or only requested pages
  – Department decides whether the digital version should be available online
Digitization Workflow

• Pre-Requisites
  – Each item has to have a barcode
  – Pagination is mandatory
  – Copyright is checked and item is under public domain
  – Orderer should get the copies he paid for before they are online
  – Department / Collection marks page that should be used for thumbnail
Digitization Workflow

• Processing
  – Scanned images are stored in a defined directory structure
    • Top level directory per item is named after the ALEPH barcode
      – Directory name parameter for X-Server request
    • File name convention for colour chart
Digitization Workflow

• Processing
  – Semi-automated workflow
    • Application developed inhouse
      – Check completeness
      – Check mandatory files (e.g. color chart)
    • Assignment of additional information
      – Thumbnail image
      – Load to DTL y/n
      – Store at Federal Computing centre y/n
  • Creation of DVD for customer
Processing

- Interface

One Object – Three Systems
Processing

• Parameters submitted by the interface are stored in a file for further automated processing

• barcode=+Z95163900;date_str=Wed Jul  6 09:58:18 2011;date_num=1309939098.07866;brz=true;digitool=true;exists_on_brz=false;data_source=vd/thumb=6;abt=MUS2;flag=2;leerseiten=9
Processing

• Application automatically submits 3 X-Server requests in the background
  – Request for Session-ID
    • not mandatory but used to lower amount of sessions
  – Request for barcode
    • Returns set number
  – Present Set
Processing

• Request for Session-ID
  – http://aleph18.onb.ac.at/X?op=login&user_name=XXXX&user_password=XX XX&library=onb01

• <login><session-id>18AB1VMBNVBBH6IRAFLG13BE6BDLMB6EKB1AN74CC2 MR29KM2R</session-id></login>

• Request work without login and password

  – Make sure to restrict X-Server access in server_ip_allowed
Processing

• Request for barcode

  – http://aleph18.onb.ac.at/X?op=find&base=onb01&code=bar&request=%2bZ43207201&session=ADYVEPY5KA4PRVDU6AF1RA3E3Y4SRI8HKG6CJTXY4DQ1L7XFUH

  • returns set number for result

    – <find><set_number>078831</set_number><no_records>000000001</no_records><no_entries>000000001</no_entries><session-id>ADYVEPY5KA4PRVDU6AF1RA3E3Y4SRI8HKG6CJTXY4DQ1L7XFUH</session-id></find>
Processing

- Present Set
  - http://aleph18.onb.ac.at/X?op=present&set_no=078831&set_entry=1&session=ADYVEPY5KA4PRVDU6AF1RA3E3Y4SRI8HKG6CJTXY4DQ1L7XFUH
Processing

- Present

```xml
<present>
<record>
<record_header>
<set_entry>000000001</set_entry>
</record_header>
<doc_number>000435100</doc_number>
<metadata><oai_marc>
<fixfield id="LDR">-----nM2.01200024------h</fixfield>
<fixfield id="FMT">MH</fixfield>
<varfield id="001" i1="-" i2="1">
<subfield label="a">AC03060708</subfield>
</varfield>
<varfield id="002" i1="a" i2="1">
<subfield label="a">20001018</subfield>
</varfield>
</oai_marc>
</metadata>
</record>
</present>
```
Processing

• Background Operations
  – XML-Output of present request is stored in image directory
    • File is named with 001-Identifier
  – Service Copies
    • Low resolution images are created automatically
    • Thumbnail is created
Processing

• Background Operations
  – Transfer of archive copies to storage at Federal Computing Center
  – Transfer of service copies to temporary storage
  – A list of all barcodes with creation date of each directory is provided
    • Used by ingest application
Ingest

- Ingest Maintenance Application
  - Small database
    - Inhouse development
      - Status of ingest
      - Guarantees:
        » No redundant ingest
        » User requested objects get an embargo (6 weeks)
        » Request from library are handled immediately
    - Contains parameters for command line ingest script
Ingest

- Command line ingest
  - Tasker.sh
    - Transformer
      - METS
    - Tasks
      - Add Metadata (DC)
        » dc:alephszncid
      - Attribute Assignment
        » Barcode in partition_b
      - Thumbnail for COMPLEX/METS
Ingest

• METS-File
  – Archive copies referred by URL
  – Low resolution service copies for physical ingest in DTL
  – Very simple logical structure (only pagination)
  – Thumbnail is object #1 in file_sec
Ingest

— METS

- <mets:fileSec>
- <mets:fileGrp USE="archive">
- <mets:file ID="FID00000001" MIMETYPE="image/jpeg" SEQ="1" GROUPID="GID00000001">
- </mets:file>
- <mets:file ID="FID00000002" MIMETYPE="image/jpeg" SEQ="2" GROUPID="GID00000002">
- </mets:file>
- <mets:file ID="FID00000003" MIMETYPE="image/jpeg" SEQ="3" GROUPID="GID00000003">
- </mets:file>
- <mets:file ID="FID00000004" MIMETYPE="image/jpeg" SEQ="4" GROUPID="GID00000004">
DigiTool -> ALEPH

- DTL-ALEPH sync
  - Add object (z403) to ALEPH
  - Get full descriptive from ALEPH to DTL
  - Replication: 4 times/day
  - Synchronisation: Only once per ALEPH library (night)
Primo

• Digital objects are available in both ALEPH and Primo
  – To avoid redundant records in Primo:
    • DTL objects linked to ALEPH are published from ALEPH only
      – No thumbnail (workaround will be presented later)
      – No fulltext (objects are images only)
Primo

• Consortia
  – Primo in consortial installation, hosted by OBV
  – Central catalog in ALEPH with downward replication to local catalogue
    • No DTL-ALEPH sync with Central catalogue
Primo

• Primo-Publishing
  – Primo publishing from ALEPH is done from central catalogue
    • Links to DTL not persistent
    • Links to DTL not available in central catalogue -> loss
    • -> Mechanism required to get persistent links into Primo
Primo

• DTL-Links in Central Catalogue
  – Requirement
    • ANL maintains re-direct to offer persistent links
    • re-direct from stable URL to DTL Delivery URL
  – Processing
    • OBV checks ALEPH z403 once/week
      – Identifies pids
      – Creates persistent URL with PID
Primo

- DTL-Links
  - OBV adds persistent URL to central catalogue records
    - DTL links published to Primo
      - Central view
      - Local views that own title
    - DTL links replicated to local ALEPH catalogues
      - Distribution to all catalogues that own title
Primo

• Thumbnails
  – Get Thumbnails from DTL
    • Thumbnails not present in ALEPH
      – Mechanism to get thumbnails to Primo required
        » How to find thumbnails for links from Primo?
        » Link in Primo points to VIEW-manifestation
        » PID of thumbnail not known in Primo
Primo

• Thumbnails
  – Mechanism
    • Script submits an sql request to DTL to identify all thumbnails
    • All thumbnails are copied to new directory with PID of VIEW manifestation as filename
Primo

• SQL

  SELECT
  a.id,
  c.root,
  b.internalpath,
  e.pid
  FROM
  D31_REP00.HDECONTROL a,
  D31_REP00.HDESTREAMREF b,
  D31_REP00.HRESTORAGE c,
  D31_REP00.HDERELATION d,
  D31_REP00.HDECONTROL e
  WHERE
  e.pid LIKE '%%%d'
  e.pid not like 'DTL%%'
  AND (a.usagetype = 'THUMBNAIL')
  AND (a.id = b.id)
  AND (c.appid = b.storageid)
  AND (d.control = a.id)
  AND (e.id = d.targetcontrol)

  ORDER BY
  a.id DESC
Primo

• Primo-Setup for thumbnail
  • Additional Data Section (local field)
    – Takes link in record to VIEW- Manifestations and extracts PID with regular expression
Primo

• Primo-Setup for thumbnail
  • Link Template
    – Replaces URL prefix
    – Adds PID to new URL that points to thumbnails
      » http://data.onb.ac.at/dtl/PID (objects)
      » http://data.onb.ac.at/dtlthm (thumbnails)
Primo

- Primo-Setup for thumbnail
- PNX-Link Section
  - Takes generated new URL

### Rule group: links_thumbnail

<table>
<thead>
<tr>
<th>Type</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>Constant</td>
</tr>
</tbody>
</table>

#### Conditions

- **Condition 1 - Logic**: True
- **Condition 1 - Source**: True
- **Field**: 655
- **Ind1**: 0
- **Ind2**: 0
- **Last**: 0
- **Step**: 0
- **Subfield**: Include
- **Success if**: Match Any

#### Transformations

- **Transformation**: Check string equals string
- **Parameter**: ONB-DTL
Primo-Examples

- Mainz Psalter

<table>
<thead>
<tr>
<th>1 Ergebnisse für ÖNB Gesamtbestand</th>
<th>sortiert nach: Relevanz ▼</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bücher</td>
<td>Bestellen</td>
</tr>
</tbody>
</table>

– [Link](http://search.obvsg.at/primo_library/libweb/action/display.do?tabs=detailsTab&doc=ONB_aleph_acc003762520&&displayMode=full&tab=default_tab&vid=ONB)
Primo-Examples

• Gutenberg Bible

<table>
<thead>
<tr>
<th>Ergebnisse für ÖNB Gesamtbestand</th>
<th>sortiert nach: Relevanz</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Bibbia</td>
<td></td>
</tr>
<tr>
<td>[Mainz] : [Druckerei der 42-zeiligen Bibel, d.i. Johann Gutenberg, Johann Fust und Peter Schöffer] [um 1454-55, nicht nach August 1456]</td>
<td></td>
</tr>
</tbody>
</table>

Prüfen Sie die Verfügbarkeit

Bücher Bestellen Standorte Details

http://search.obvsg.at/primo_library/libweb/action/search.do?tab=onb_digital&indx=1&vl(freeText0)=AC07988562&vid=ONB&fn=search
Conclusion

• Technologies
  – No sophisticated new technologies
  – Existing tools
  – Some PERL scripting

• Project
  – Complexity
    • 3 applications
    • Local vs. Central
Conclusion

• Project
  – Cooperation
    • Implementation not possible without cooperation with OBV
    • Seamless integration of features
    • Win2Win
      – ANL gets required functionality
      – Consortia gets links to objects to distribute in all participating systems
Outlook

- DTL-objects with Access Rights
  - Primo RTA
    - DTL objects harvested from ALEPH don’t have information whether objects are not free
  - Workaround
    - If objects is protected information is copied as string to DTL label
    - Label is replicated to ALEPH und harvested into Primo
Outlook

• Orders
  – Integrated Order forms from Primo missing
    • Integration of non Ex Libris Systems
    • -> SFX (?)

• ABO (Austrian Books Online)
  – Google Books digitization
  – Not handled with DigiTool
  – Access via Primo
    • Fulltext
THANK YOU!

dirk.willinghoefer@onb.ac.at