Scriptorium, a retro-cataloguing tool to easily and quickly encode older book items

François Renaville, University of Liege Library
with Sylvain Danhieux, University of Liege Library

Download: http://hdl.handle.net/2268/171960
Abstract

The University of Liège Library’s books collection is composed of some 2,000,000 print volumes, of which only 60% are catalogued. Most of the uncatalogued books have been published before 1970 and cataloguing these according usual standards and norms would certainly take decades. To decrease the cost of treatments and increase the number of catalogued volumes, the Library developed a light PHP/MySQL application, Scriptorium, that enables non-catalogers (mostly students) to quickly encode the books (ca 3 min. per item) by providing the most essential information. References are then daily exported in MarcXML to the ILS. Scriptorium has been developed to permit to easily create new independent instances for different parallel retro-cataloguing projects and also to be used by other libraries.
University of Liege Library

• University founded in 1817
• 5 main libraries:
  • Arts & Humanities Library
  • Agronomy Library
  • Law & Social Sciences Library
  • Life Sciences Library
  • Science & Technology Library
• 18 library facilities (branches)
• + 1 Remote Storage

4 campuses:
  ▶ Liège Sart-Tilman (main campus)
  ▶ Liège City centre
  ▶ Gembloux
  ▶ Arlon
Library Print Collections

- More than 2,000,000 books
  - 1,200,000 catalogued books
  - 800,000 uncatalogued books

Also:
- 50,000 maps
- 30,000 dissertations
- 6,500 manuscripts
- 565 incunables
- ...

- Automation started in the 1960s
- But only in the 1990s for some HS branches (big collections)
- Many books located in the Remote Storage
  - +/- 300,000 catalogued books
  - +/- 500,000 uncatalogued books
RFID as a catalyst

• In Spring 2014, RFID project for a library branch
  • History Library: more than 80,000 books
  • Around 10,000 not yet catalogued (12.5%)
    • Documents related to the national or local history
    • Produced between 1800 and ca 1970
    • Items to keep (not for pulp!), even if not on loan for years!
  • Types
    • Published books
    • Reports
    • Master dissertations
    • Offprints

• For the RFID project, all items had to be catalogued (organisational reasons)
Aleph Quick Catalog function

- Not very happy with Aleph Quick Catalog function
  - Two tabs (doc info + item info) -> no global view
  - Same configuration for all -> not very flexible, not context sensitive
  - Few possibilities for default values (only for item info)
  - Limited number of ccu (40 in our case)
  - Drop down lists may have irrelevant content
  - No duplicate function
  - ...
Homemade Quick Catalog Web Application

- Requested specifications
  - Flexible and context sensitive
  - For an unlimited number of users
  - Easy to use, especially made for non-cataloguers
    - Students
    - Low qualified library agents
    - Non-librarian staff (accountants, secretaries...)
    - Trainee librarians
    - Heads of sections, heads of services...
  - Not only for BIB records and items, but also able to provide info in HOL records
  - To be used for any retro-cataloguing project to come (eg. Remote Storage): each retro-cataloguing project has its own instance
  - Also available for non ULg libraries: easy to deploy in other contexts
- Developed with PHP/MySQL
Leitmotiv for the Application

• For those older and un-catalogued collections:
  • **VISIBILITY**
    • Better in the ILS than not in the ILS
  • **QUANTITY** over Quality
    • Better in the ILS, even with poor metadata, than not in the ILS
    • As many items catalogued as possible
  • **SPEED**
    • ASAP added into the ILS
Where anyone (employed students, library staff, low qualified agents...) could encode older uncatalogued books....
4 Roles defined

**ENCODER** can
- create new items
- delete/modify items he/she has created
- define default values for encoding

**MANAGER** can
- Same as ENCODER
- + delete/modify any items created by any encoder
- + export records

**ADMIN** can
- Same as MANAGER
- + manage the accounts
- + manage the encoding form (metadata structure, mandatory or not, mapping with Marc fields)

**SUPERADMIN** can
- Same as ADMIN
- + manage ADMIN accounts
- + create instances
- + manage backup
- + manage export rules
Scriptorium, a retro-cataloguing tool...
### How the Admin Can Edit/Create a Field

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>form_field_id</code></td>
<td>0</td>
</tr>
<tr>
<td><code>label</code></td>
<td>Language</td>
</tr>
<tr>
<td><code>type_id</code></td>
<td>Select</td>
</tr>
<tr>
<td><code>required</code></td>
<td>✓</td>
</tr>
<tr>
<td><code>read_only</code></td>
<td>✗</td>
</tr>
<tr>
<td><code>default_value</code></td>
<td></td>
</tr>
<tr>
<td><code>authority_code</code></td>
<td>1</td>
</tr>
<tr>
<td><code>tag</code></td>
<td>041</td>
</tr>
<tr>
<td><code>indicuteur_1</code></td>
<td></td>
</tr>
<tr>
<td><code>indicuteur_2</code></td>
<td></td>
</tr>
<tr>
<td><code>souschomp</code></td>
<td>a</td>
</tr>
<tr>
<td><code>option</code></td>
<td>✓</td>
</tr>
</tbody>
</table>

**Option:** Allows the encoder to define default content within a session

**Other types:**
- Textbox
- Textarea
- Checkbox
- Radio
- Hidden
### authority_code

<table>
<thead>
<tr>
<th>Drag</th>
<th>Id</th>
<th>Label</th>
<th>ID</th>
<th>Value</th>
<th>Displayed Value</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>✗</td>
<td>29</td>
<td>Language</td>
<td>1</td>
<td>fre</td>
<td>Français</td>
<td></td>
</tr>
<tr>
<td>✗</td>
<td>30</td>
<td>Language</td>
<td>1</td>
<td>eng</td>
<td>English</td>
<td></td>
</tr>
<tr>
<td>✗</td>
<td>31</td>
<td>Language</td>
<td>1</td>
<td>ger</td>
<td>Deutsch</td>
<td></td>
</tr>
<tr>
<td>✗</td>
<td>32</td>
<td>Language</td>
<td>1</td>
<td>dut</td>
<td>Nederlands</td>
<td></td>
</tr>
<tr>
<td>✗</td>
<td>33</td>
<td>Language</td>
<td>1</td>
<td>spa</td>
<td>Español</td>
<td></td>
</tr>
<tr>
<td>✗</td>
<td>34</td>
<td>Language</td>
<td>1</td>
<td>ita</td>
<td>italiano</td>
<td></td>
</tr>
<tr>
<td>✗</td>
<td>35</td>
<td>Language</td>
<td>1</td>
<td>lat</td>
<td>Latin</td>
<td></td>
</tr>
<tr>
<td>✗</td>
<td>36</td>
<td>Language</td>
<td>1</td>
<td>und</td>
<td>Undefined</td>
<td></td>
</tr>
</tbody>
</table>

### options

#### Options

<table>
<thead>
<tr>
<th>Language</th>
<th>Collection</th>
<th>Material Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please Select One</td>
<td>Please Select One</td>
<td></td>
</tr>
<tr>
<td>Français</td>
<td>English</td>
<td>Deutsch</td>
</tr>
<tr>
<td>Español</td>
<td>Nederlands</td>
<td>Español</td>
</tr>
<tr>
<td>Italiano</td>
<td>Latin</td>
<td>Undefined</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Submit</th>
<th>Cancel</th>
</tr>
</thead>
</table>
Exports

- Automatic daily export from Scriptorium to Aleph
- Manual export is also possible (csv or MarcXML)
<record>
  <leader>^^^^nam^a22^^^^za^4500</leader>
  <controlfield tag="001">00011</controlfield>
  <controlfield tag="008">^^^^s^^^^^^^^^^^^^|||^^^^|||||eng^d</controlfield>
  <datafield tag="245" ind1="0" ind2=" ">
    <subfield code="a">The Art and Practice of Diplomacy</subfield>
  </datafield>
  <datafield tag="720" ind1="1" ind2=" ">
    <subfield code="a">Webster, Charles</subfield>
  </datafield>
  <datafield tag="041" ind1=" " ind2=" ">
    <subfield code="a">eng</subfield>
  </datafield>
  <datafield tag="260" ind1=" " ind2=" ">
    <subfield code="c">1961</subfield>
  </datafield>
  <datafield tag="910" ind1=" " ind2=" ">
    <subfield code="a">gamassange</subfield>
  </datafield>
  <datafield tag="910" ind1=" " ind2=" ">
    <subfield code="b">20131107_1005</subfield>
  </datafield>
  <datafield tag="952" ind1=" " ind2=" ">
    <subfield code="1">BGP06</subfield>
    <subfield code="2">PH002</subfield>
    <subfield code="5">13BGP21395</subfield>
    <subfield code="j">HC/A.0020 (-)</subfield>
    <subfield code="m">BOOK</subfield>
  </datafield>
</record>

Year in 008/7-10 and 245 2nd ind. are added with Aleph programs (tab_fix)

910 = admin values (encoder account + creation date)
Hidden in the New Item form

In the export rules, any content that should be in a 952 is concatenated into one single 952.
Aleph Scripts

- BIB records, holdings and items created with Aleph services
- No deduplication verification with existing Aleph records
- job_list:
  - p_file_02
    USM01, script_%DATE.xml, script_%DATE_alseq, 06,
  - p_manage_18
    USM01, script_%DATE_alseq, script_%DATE_alseq.reject, script_%DATE_alseq.doc_log, NEW, SCR, FULL, APP, M, SCRIPTORIU,
  - p_ret_03
    USM01, script_%DATE_man50, WCDA=SCRIPTORIU AND WCDA=%DATE,
  - p_manage_50
    USM01, script_%DATE_man50, 000000000, 999999999, USM51, USM61, 952##, tab_hol_item_create_scriptorium, A, A, N, SCRIP TORIU, 30, SCRIPTORIU, 30, Y, N,
  - p_manage_37
    USM01, DOC_LIST, 001914034, 999999999, script_%DATE_man37, SCR12, Y,
USM01,script_%DATE.xml,script_%DATE_alseq,06,

Where:
• USM01 = Bibliographic Base
• script_%DATE.xml = input file
  (%DATE creates date relative to the current date)
• script_%DATE_alseq = Output File
• 06 = MarcXML Format
Where:

- **USM01** = Bibliographic Base
- **script_%DATE_alseq** = Input File (from file_02)
- **script_%DATE_alseq.reject** = Output File for Rejected Records
- **script_%DATE_alseq.doc_log** = Output File for Logging System Numbers
- **NEW** = Add new records
- **SCR** = Fix Routine
- **FULL** = Full Indexing
- **M** = Multi-user
- **SCRIPTORIU** = Cataloger Name

```plaintext
tab_fix
SCR fix_doc_do_file_08  fix_scriptorium
SCR fix_doc_createFmt
SCR fix_doc_001
SCR fix_doc_005
SCR fix_doc_tag_008_openDate
SCR fix_doc_tag_008
SCR fix_doc_punctuation_usm
SCR fix_doc_non_filingInd
SCR fix_doc_delete_empty
SCR fix_doc_sort
```
USM01,script_%DATE_man50,WCDA=SCRIPTORIU AND WCDA=%DATE,

Where:
- USM01 = Bibliographic Base
- script_%DATE_man50 = Output File to be used with man-50
- WCDA=SCRIPTORIU AND WCDA=%DATE = CCL search phrase
p_manage_50

USM01, script_%DATE_man50, 000000000, 999999999, USM51, USM61, 952##, tab_hol_item_create_scriptorium, A, A, N, SCPICTORIU, 30, SCPICTORIU, 30, Y, N,

Where:

• USM01 = Bibliographic Base
• script_%DATE_man50 = Input File (from ret-03)
• USM51 = Admin base
• USM61 = HOL base
• 952## = Main Field
• tab_hol_item_create_scriptorium = Item Creation Table

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>!</td>
<td>!</td>
<td>!</td>
<td>!</td>
<td>!</td>
<td>!</td>
<td>!</td>
</tr>
<tr>
<td>HOL-040</td>
<td>Y</td>
<td>a</td>
<td>040</td>
<td>d</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CALL-NO-TYPE</td>
<td>a</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOL-852</td>
<td>b</td>
<td>Y</td>
<td>a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOL-852</td>
<td>1</td>
<td>Y</td>
<td>b</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOL-852</td>
<td>2</td>
<td>Y</td>
<td>c</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOL-852</td>
<td>j</td>
<td>Y</td>
<td>j</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOL-852</td>
<td>q</td>
<td>Y</td>
<td>q</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOL-852</td>
<td>x</td>
<td>Y</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z30-BARCODE</td>
<td>5</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z30-ITEM-STATUS</td>
<td>f</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z30-MATERIAL</td>
<td>m</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z30-CIRCULATION</td>
<td>N</td>
<td>590</td>
<td>b</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Scriptorium, a retro-cataloguing tool...
USM01, DOC_LIST,, 001914034, 999999999, script_%DATE_man37, SCRI2, Y,,

Where:

• USM01 = Bibliographic Base
• DOC_LIST = Input File Type
• 001914034 = From Document Number <001914034> (must be updated regularly to avoid too large ranges!!)
• 999999999 = To Document Number
• script_%DATE_man37 = Output File
• SCRI2 = Fix Routine

```
tab fix  
SCRI2 fix_doc_do_file_08  fix_scriptorium_delete  
(deletes fields become useless)```
Scriptorium, a retro-cataloguing tool...

+ Circulation Note added with `tab_hol_item_create`

⇒ BIB can be recatalogued after loan
Scriptorium, a retro-cataloguing tool...
Results

• Deployed in 3 library branches
  • -> 3 instances
  • 1st Instance for the History Section
  • 10,000 books to encode
    • Started in Mid-November 2013, finished less than 5 months later
    • Ready for the RFID Project

• For the 3 instances, more than 20,600 books added in the ILS in 9 months time:
  • No encoding everyday!
  • Most of them published between 1950 and 1979
  • Some were also published in the 2000s!!
  • 1,882 have got an ISBN
    • Most recent books are properly recatalogued!
  • Better to check every morning if night import ran successfully
Publication Years of All Scriptorium Items

<table>
<thead>
<tr>
<th>Decade</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1899-1909</td>
<td>1,521</td>
</tr>
<tr>
<td>1900-1909</td>
<td>696</td>
</tr>
<tr>
<td>1910-1919</td>
<td>657</td>
</tr>
<tr>
<td>1920-1929</td>
<td>1,029</td>
</tr>
<tr>
<td>1930-1939</td>
<td>1,371</td>
</tr>
<tr>
<td>1940-1949</td>
<td>1,491</td>
</tr>
<tr>
<td>1950-1959</td>
<td>2,865</td>
</tr>
<tr>
<td>1960-1969</td>
<td>4,839</td>
</tr>
<tr>
<td>1970-1979</td>
<td>3,600</td>
</tr>
<tr>
<td>1980-1989</td>
<td>879</td>
</tr>
<tr>
<td>1990-1999</td>
<td>298</td>
</tr>
<tr>
<td>2000-2016</td>
<td>16</td>
</tr>
</tbody>
</table>

Scriptorium, a retro-cataloguing tool...
How many minutes per record?

Number of Minutes per Record, including Manipulations
(for one of the Instances)

- Stud1: 3.5 min.
- Staff1: 4.6 min.
- Staff2: 3.3 min.
- Lib1: 4.6 min.
- Lib2: 3.3 min.

**Stud** = Students; **Staff** = (Low Qualified) Library Staff; **Lib** = Librarians/Cataloguers
Conclusions

• Probably not achieved with Aleph Quick Catalog

• Project well accepted by the librarians
  • Even by catalog specialists
  • Opportunity to have less uncatalogued items
  • Some other library branches would like an instance

• Positive side effects
  • Team spirit: Collaborations between librarians have become more intense (some from other libraries came to help)
  • Low qualified colleagues happy to help -> work acknowledged

• Additional jobs to enrich the metadata of the most recent books?

• Open to create instances for any external partner/library