Renovating bibliographic description:

Bibliothèque nationale de France response

By

Christian Lupovici

Rationale for a new bibliographic system

- **Web technology has become a standard and has induced:**
  - New interface design and new users behaviours in searching
  - Simple searching interface and hidden sophisticated tools and structured data
  - Navigation and browsing
  - Demand to access not only the bibliographic record but directly the document (library materials are not anymore only analogue, but digital)
  - Emergence of new types of services, not only finding documents, but also information around, and user contribution
  - "Copernican revolution" : the user is at the centre of the world, not anymore the library !

- **Obsolescence of BnF information system**
  - No more enhancement of the today catalogue possible
  - Need to build an up-to-date technology information system
  - New objective : to be 'in the flow' of users' practices on the Web

Cataloguing rules in France

- **1st period (1971-2003): national work conducted in parallel with international standardization**
  - ➢ Under the ISBD framework, national cataloguing rules
    - ISBD was in construction and lack of a standard on access points
    - Need for an authority file for Personal names, corporate names and titles
- **2nd period (2003-2008): BnF working first for international cataloguing rules**
  - ➢ International meeting of experts for an International cataloguing code (IME-ICC): BnF participation
  - ➢ BnF participation in the work on *Resource Description and Access* (RDA) to make RDA be acceptable by BnF and the French library community

The use of FRBR

- **The special case of FRBR**
  - ➤ BnF have worked since the beginning on FRBR
    - Patrick Leboeuf, one of the world specialists of the model
  - ➤ A major contribution on bibliographic data modeling
    - Up to now only an intellectual issue (no implementation in BnF information system possible)
    - Tomorrow: project of FRBR implementation for display
- **BnF working on FRAR**
- **A large public access to data through the biggest search engines**
  - ➤ Generation of landing pages at the Work level
    - Suppose working on bibliographic data and on the Title authority file

From bibliographic data to metadata

- **Metadata: not only bibliographic record but all data elements useful to manage digital material**
  - A need for managing digital documents
    - BnF has started digitization of its collections in 1992 for Gallica its digital library (80 000 books in image mode)
    - BnF has an ongoing mass digitization program to feed Gallica2 (full text searching and up-to-date web interface)
    - BnF is building a long term preservation repository (SPAR)

- **Using different metadata schemas as appropriate**
  - METS, MPEG21/DIDL, PREMIS, EAD …

- **Need for a convergence between the catalogue and digital material format metadata schemas**

Data flows: changes to come

- **BnF information system is local and closed**
  - Designed to be at the origin of all data production
    - No import of authority data possible
    - Not designed for descriptive data import
    - A heavy input process: a traditional cataloguer work imitation

- **To make a BnF information system open to the world**
  - To be able to import and integrate different data sources in a single workflow
    - From publishers, printers, booksellers, other providers (OCLC, SUDOC, Wikipedia…)
  - To make the data usable outside the catalogue
    - 1st step : OAI-PMH (already set up)
    - Authority records as resources to be exported and used in other environments than the catalogue (Wikipedia, Rights management societies, semantic web)

Formats evolution

- **The ISO 2709 period**
  - MARC format, one of the first formats using variable length fields.
    - BnF adopted Intermarc for its catalogue (with some other French, Belgium and Suisse libraries at the beginning of library automation)
  - Old format, not adapted anymore to new information systems, new products or services (like full text indexing)

- **The XML period**
  - Move from Marc to XML
    - MARCXML(Marc21); MarcXchange (ISO more generic); "InterXmarc" (BnF) to be able to process the data with XML tools
  - Genuine XML schemas
    - XML/EAD, Dublin core, MPEG 21/DIDL and MODS, METS…

- **Tomorrow : the RDF period**
  - More suitable than XML for flexible data use and re-use
  - Allows the coexistence and interoperability of a variety of different formats (and historical layers of heterogeneous data: important for library catalogues)
    - Ex. : BnF already uses RDF for digital preservation in SPAR

Architecture issues

- **Marc formats processed in relational databases**
  - ➤ The data model is rigid, becoming too complex and not able to be processed in a modern environment
  - ➤ We find the limits of having one integrated format for a great diversity of materials
  - ➤ Insufficient possible evolution of the service for the end-user
- **XML schemas processed in XML databases**
  - ➤ XML gives a choice of suitable formats for special material (i.e. EAD for manuscripts collections)
  - ➤ Transformations are easier from XML (InterXmarc to DC, to MPEG21 DIDL, …) so is the re-use of the data (OAI-PMH)
  - ➤ XML can be processed by a full text search engine -> providing a "one box" search ("*mots notices*"= any word in a record)

Architecture issues for the near future

- **Toward RDF databases**
  - A data model that will allow multiple views on the collections and search & browse interfaces adapted to a variety of user expectations (faculties, library professional staff and the public at large…)
  - Libraries should contribute to the "Linking Open Data" initiative
    - Making library structured data available on the semantic Web
    - But valuable only if it comes with an international participation of the library community
- **International cooperation under the auspices of OCLC ?**
  - A question back to OCLC

Thank you for your attention

Time for questions