Cataloguing and Indexing Trends

Presentation for Discussion at the National Library of Jamaica’s Seminar:

Building the National Union Catalogue: Challenges and Issues
April 21, 2009

Presented by Hyacinth Brown
Presentation will briefly examine:

- Characteristics of IT environment and impact on publishing, libraries and users of information
- Developments in Cataloguing and Indexing
- Implications

Building the Jamaica Union Catalogue
Cataloguing and Indexing Trends

Characteristics of ICT

- Powerful computers – memory chip capable of more than a trillion calculations per second (teraflops)
- Storage devices with large storage capacity
- Complex and variety of computer networks with Internetworking representing the interconnection of many sub-networks.
- Internet widely available commercially, in public libraries, schools and other academic institutions, in the home.
Cataloguing and Indexing Trends

Impact of IT

Will be reviewed under three headings:

- Publishing
- Libraries
- Users
Cataloguing and Indexing Trends

Impact of IT: Publishing & Publications

Vast amount of information published in print & electronic formats whether textual, a/v, images, music, moving images, etc.

• Huge amount of information published on the Internet
  – in any field,
  – published by anyone from government to prep school child
  – valuable and otherwise

• Organizations/corporate bodies, groups and individuals publish extensively on the Internet
Cataloguing and Indexing Trends

Impact of IT: Publishing & Publications

• Faster publishing cycle - from creator to user virtually instantly
• Publishing is easy and not as costly as with traditional publishing
• Most textual, graphics, photographic and musical publications born digital
  – Some published with indexing data, i.e., metadata
Indexing of publications

- Metadata indexing provided by libraries through their OPACs, for access to their print and electronic collections & some Internet resources
- Metadata indexing of Internet resources done by some publishers and provides for more reliable results in response to a search
  - Number of metadata schema have been developed for indexing digitized resources
- Automatic indexing of Internet resources using Web Crawlers done by Search engines – provides for “easy” searching
  - Search engine indexing lacks control of terms used. So often recall very high but relevance low
- Some search engines mine metadata and use it in their indexes
Cataloguing and Indexing Trends

Impact of IT: Users

- Organizations, corporate bodies, individuals use the Internet for publishing and finding information
- Information used not necessarily accessed via library OPACs which can be accessed via search engines
- Users prefer Google-like search capabilities
- Users expect items found to be available immediately – easy to get a copy whether purchase or electronic digitized or subscription – quick access
Impact of IT: Libraries

- Automated processes for many library functions
- OPACS
- Still committed to bibliographic control but unable to organize most of what is on the Internet
- Use of bibliographic utilities for cataloguing just to keep abreast of non-Internet publications
- Digitized collections in all types of libraries
- Provide access to union catalogues
- Provide electronic databases for journal articles and technical reports
If impact of IT means

- More publications, fast publication cycle; but library unable to organize them
- Many available on the Internet and in great demand,
  - Various types,
  - Directed at great variety of users,
- Most born digital whether published as printed or electronic resources
- Many user communities have developed own metadata schema for publishing on the Internet, so some publications born with metadata – not AACR2 based
- Many Internet publications do not have metadata so indexed by search engines resulting in poor quality of indexing,
- Users prefer searching Google to searching catalogues/libraries
- Users expect quick access to resources
If also accept that for retrieving resources whether in library or on the Internet:
Descriptive elements need to be identified
Need authority control of subject terms, names of persons, organizations etc.
Library community has the expertise, skills and has developed the body of knowledge for organizing information but cannot do it alone.
Others involved in organizing need the expertise and skills – Is AACR2 answer for them? Are subject heading lists?
Then how should an entity interested in organization of information for easy access respond …

Equip creators of publications with tools and skills to catalogue or add metadata when publishing as this will allow for better searching and retrieval resulting in higher relevance of items found in response to a search query.

This is in essence what the library community is doing.
Will examine four developments:

- Dublin Core
- FRBR
- RDA
- FAST
The Dublin Core Metadata Element Set or DC came out of the library field in response to large amount of resources on Internet and speed with which these resources were being published proving to be a significant challenge to Bibliographic Control.

Developed by OCLC
Objective was to have a simple, easily understood, international standard for descriptive elements traditionally used in libraries. Could be used by any creator or publisher in describing their resources. This would help to improve “identification and discovery” as Search engines could use these data or metadata rather than using automatic indexing techniques in creating their indexes.
Cataloguing and Indexing Trends

Developments in Cataloguing & Indexing

Dublin Core

Title
Subject & Keywords
Description - an account of the content
Resource type - image, text, sound, software etc.
Source - a resource from which the present resource is derived.
Relation - e.g. may be part of a whole item
Coverage - could be period or place
Creator
Publisher
Contributor - An entity responsible for making contributions to the content of the resource
Rights - e.g. copyright and property rights
Date
Format
Identifier e.g. URL
Language
Audience


5/21/2009

Building the Jamaica Union Catalogue
IFLA’s Functional Requirements for Bibliographic Records – FRBR - 1998

Purpose was to “produce a framework that would provide a clear, precisely stated, commonly shared understanding of what it is that the bibliographic record aims to provide information about, and what it is that we expect the record to achieve in terms of answering user needs.”
FRBR conceptualizes three groups of entities:

- **Group 1** consists of the products of intellectual or artistic endeavor (e.g., publications).
- **Group 2** comprises those entities responsible for intellectual or artistic content (a person or corporate body).
- **Group 3** includes the entities that serve as subjects of intellectual or artistic endeavor (concept, object, event, and place).
Developments in Cataloguing & Indexing

FRBR

The internal subdivision of Group One entities is important as well. FRBR specifies that intellectual or artistic products include the following types of entities:

- **the work**, a distinct intellectual or artistic creation
- **the expression**, the intellectual or artistic realization of a work
- **the manifestation**, the physical embodiment of an expression of a work
- **the item**, a single exemplar of a manifestation.
FRBR also specifies particular relationships between classes of Group One entities:

- a **work** is realized through one or more **expressions**
  - each of which is embodied in one or more **manifestations**
- each of which is exemplified by one or more **items**

Source: http://www.oclc.org/research/projects/frbr/
Cataloguing and Indexing Trends

Developments in Cataloguing & Indexing

FRBR

Will affect:

Look of search results from a catalogue

Approach to cataloguing e.g. rather than begin with what is ME, will ask is this work already in the catalogue?
Cataloguing and Indexing Trends

Developments in Cataloguing & Indexing

FRBR

Search in OCLC’s WorldCat
Cataloguing and Indexing Trends

Developments in Cataloguing & Indexing

FRBR
Search in OCLC’s WorldCat
Cataloguing and Indexing Trends

Developments in Cataloguing & Indexing

FRBR

Same search in British Library’s Integrated Catalogue
Developments in Cataloguing & Indexing

RDA: Resource Description and Access

Need for a new code

1. Opportunity to establish it as a content standard for resource description for any metadata scheme. Content vs. carrier.
2. Opportunity to recognize and include other metadata user communities – not only skilled cataloguers
3. Need a code that is more consistent across the various types of content and media, showing commonalities of various types of resources
Need for a new code

3. The emergence of electronic resources, particularly the Internet - along with the proliferation of web-pages - has provided information formats which challenge the library’s quest for bibliographic control.

5. It is an opportunity to simplify cataloguing code. Rules have become increasingly complex to accommodate formats, exceptions and conditions.
Need for a new code
6. Rules need to be independent of any communication format and provide a true content standard and so can be used with any emerging metadata standard.
7. The rules based on physical objects rather than Internet resources which have become the format of choice for students and others. See lecture on cataloguing Internet resources.
Differences between AACR2 and RDA

1. RDA based on conceptual model of FRBR so will have description of relationships; will enable better displays in catalogues by clustering related items - e.g. translations, abridgements, different physical formats.
2. Designed for Internet resources as an online tool.
3. Content focused on display as well.
4. Single set of general rules for all types of materials rather than each being dealt with in a different chapter, thus eliminating redundancy. Rules grouped by bibliographic content rather than by format.
Differences between RDA and AACR2

5. Authority control given more emphasis and constitutes section two

6. Change in concepts
   - Main entry concept not used – use primary access point. The concept of main entry as used in a card catalogue is no longer applicable in online catalogues. Nevertheless, there is still a need to choose a primary access point to create bibliographic citations, and to collocate works and expressions in the online catalogue. RDA will provide instructions on choosing this primary access point.
   - Added entries or references called variant access points
Differences between RDA and AACR2

7. Change in use of abbreviations - Abbreviations replaced by spelt out words or phrases which may be more accurate – e.g.: s.l. replaced by place not recorded or place not known and s.n. by publisher unknown

8. Publisher instructions inserted before those for place

9. Statement of responsibility no longer restricted by ‘rule of three’. Optional – may add as many as want to

10. Treatment of parts of the Bible for uniform headings, NT and OT no longer required. E.g. Bible. Genesis NOT Bible. OT. Genesis
Cataloguing and Indexing Trends

Developments in Cataloguing & Indexing

**RDA**

Differences between RDA and AACR2

11. Change in prescriptive approach
   More emphasis being placed on cataloguer judgment

12. More logically structured and cataloguer accesses information when needed

13. Designed for use by other metadata users – encouraging a common content standard so that there can be improvement in cross searching databases and repositories. Better retrieval.
RDA timeline (subject to change):

- March 2009 – JSC and Committee of Principals meet in Chicago. JSC finalises review of comments received.
- Third quarter calendar 2009 – RDA is released.
- Last quarter calendar 2009–early 2010 – CoP national libraries evaluate RDA prior to implementation.
The Joint Steering Committee for Development of RDA (JSC), working with the RDA Editor, is responsible for developing RDA. The JSC consists of representatives from six major Anglo-American cataloguing communities. These include the American Library Association (ALA), the Australian Committee on Cataloguing (ACOC), the British Library (BL), the Canadian Committee on Cataloguing (CCC), the Chartered Institute of Library and Information Professionals (CILIP), and the Library of Congress (LC). The JSC and RDA Editor are assisted by the JSC Secretary and the RDA Project Manager.

Source: http://www.collectionscanada.gc.ca/jsc/rdafaq.html#2
Facetted Application of Subject Terminology - FAST

1. FAST headings were recently developed by OCLC from the LCSH list.

2. It is a post-coordinated faceted vocabulary for an online environment.
Facetted Application of Subject Terminology - FAST

“Recent trends, driven to a large extent by the rapid growth of the Web, are forcing changes in bibliographic control systems to make them easier to use, understand, and apply, and subject headings are no exception. The purpose of adapting the LCSH with a simplified syntax to create FAST is to retain the very rich vocabulary of LCSH while making the schema easier to understand, control, apply, and use. The schema maintains upward compatibility with LCSH, and any valid set of LC subject headings can be converted to FAST headings”

Developments in Cataloguing & Indexing

Merits are that:
• Usable by people with minimal training and experience,
• Enables a broad range of users to assign subject terminology to Web resources,
• Compatible with automated authority control,
• Compatible with use as embedded metadata,
• Is a post-coordinate system in an online environment
1. Appreciate that body of knowledge underpinning cataloguing practices:
   • has wider application than in the library field
   • Is very relevant to IT environment

E.g. principles of user needs consideration, uniformity, consistency and uniqueness in headings, use of current terminology - these make difference between ‘good’ and ‘bad’ search results

So should accept responsibility that comes with the above.
2. The surrogate record is of even more importance in environment of vast number of publications from which users have to select the few relevant ones.

3. Many options in providing descriptive data – i.e. many metadata schema to choose from including RDA but in cooperative environment standardization important.
4. Need to become familiar with more than one metadata schema

5. The Internet resources and digitized resources are the future default materials so need to adapt to the new tools for cataloguing these.
Conclusion

Developments in cataloguing and indexing, in response to the IT environment, will help to ensure that organization of the vast number of publications is achieved.

The cataloguer’s role will increasingly involve equipping creators and publishers to index their resources for efficient and effective retrieval.
THANK YOU FOR YOUR KIND ATTENTION!