

RDA and library systems

Gordon Dunsire

Abstract

The paper explains the background to the development of Resource Description and Access, a metadata content standard intended for international use by a wide range of metadata communities. The paper discusses the implications of RDA for library systems.

Introduction

RDA is the acronym for Resource Description and Access, a new standard for the content of metadata used to support the discovery, identification and employment of information resources (1). The standard is designed for the digital environment and is intended for multinational use.

It is aimed at all end-users who need to find, identify, select, obtain, or use information, and all those who need to manage and organize information resources in a professional capacity. RDA deals only with metadata content, so it is independent of any specific technical storage or communication format.

RDA is still in development. It is approximately half-way through its development schedule, and is due for publication early in 2009.

Background

RDA is based on the Anglo-American Cataloguing Rules (AACR), which is the most-used standard for bibliographic metadata content in the world.

AACR itself is based on a long pedigree of bibliographic content rules and guidelines. The first of these emerged in 1841 in the form of rules for the British Museum's library catalogue. The first American library cataloguing rules were published in 1876. From 1902 to 1949 the United Kingdom (UK) and United States of America developed separate sets of rules. The stimulus to consider developing common rules came in 1961 with the formulation of the principles of cataloguing and its endorsement by the International Federation of Library Associations (IFLA). National cataloguing agencies from the UK and North America collaborated to produce the Anglo-American Cataloguing Rules in 1967. Not all differences could be reconciled, however, and this first edition of AACR was published in distinct UK and North American versions. Pressure to eliminate those remaining differences continued with the development of a more general content standard, the International Standard Bibliographic Description, by IFLA in 1969. As a result, the second edition of AACR (AACR2) was published in 1978 in a single common version. Recognition of

international perspectives on cataloguing standards has thus been a significant influence on the evolution of AACR.

AACR has continued to develop since 1978, not least to accommodate the new forms of information resource emerging from the Internet and World-Wide Web. Following 20 years of incremental development, an international conference of experts was held in Toronto, Canada in 1997 to consider the principles and future development of AACR.

In 1998 IFLA published Functional Requirements for Bibliographic Records, an entity-relationship model intended to support the design of information retrieval and resource discovery services for end-users (2). Further IFLA activity in this area has continued more recently with a series of meetings held in different locations worldwide from 2004 to 2007 with the aim of updating and re-affirming the principles of cataloguing endorsed in Paris in 1961.

The environment within which cataloguing takes place has undergone significant change since the publication of AACR2 in 1978. Metadata are required to describe a wider range of information carriers with a greater depth and complexity of content ranging from, for example, collections of digital information resources to multi-component learning packages. Intranets and Internet allow a greater variety of personnel to contribute to the content of metadata, from authors supplying title, summary and subject analysis through administrators determining identifiers and access conditions to cataloguers applying access points and assuring quality. The issues associated with these changes in a further education environment are discussed in the toolkit created by the Managing digital assets in tertiary education project (3). Computer-generated metadata is also of increasing significance. Rules and guidelines must be clear, unambiguous, consistent, and detailed and structured enough to be applied effectively across such a range of skills and abilities. Metadata content must also be structured enough to be expressed in modern technical storage and communication formats for utilisation by end-user services. There has been a proliferation of such formats for resource description and access in recent years, with the various flavours of Machine-Readable Cataloging (MARC) being joined by Dublin Core (DC), the Institution of Electrical and Electronics Engineers Learning Object Metadata (IEEE-LOM), and many others. RDA is discussing its potential relationship with these formats with representatives of both the DC and IEEE-LOM communities.

Impact on library systems

The most immediate and obvious impact that RDA will have on library management systems (LMS) is improved integration of metadata content rules and guidance with cataloguing modules. RDA will be published as an online product offering direct access to individual rules, glossary terms and other specific content as well as structured navigation across and within its component parts. LMS vendors are being kept informed of RDA so that they can develop their cataloguing workflows, input and amendment templates, and context-sensitive help services to utilise the functionality of the RDA product. A proof-of-concept demonstrator of how this might work is available (4).

The product should allow the addition of local notes and examples, and the concealment of unused options. RDA contains options and detail that will not be required by every library community. It is quite possible that some communities will take a rule-inheritance approach by defining which general options in RDA are to be used, and then allowing sub-communities to further define options. For example, a national cataloguing agency might determine the overall policy for that country's adoption of RDA, including language and script for recording content as well as, say, method for expressing relationships between metadata entities. Higher education libraries within that country might then further decide on what level of granularity to describe resources, and the audiovisual department within a specific university library would then customise their view for RDA to remove all rules for non-audiovisual resources.

A corollary of integrating RDA into a cataloguing help and support system is to consider using it for helping advanced end-users of the resultant metadata. RDA is eliminating much of the jargon and arcane terminology used in AACR, such as abbreviated Latin phrases like "et al." and "s.n.". The content of RDA is written in plain English and abbreviations are generally avoided, and the glossary will clearly define terms with special meanings. Expert users of online catalogues might well benefit from seeing, for example, why content is transcribed in certain ways or the relationship between the resource and its description.

There is likely to be little direct impact on the functionality and design of general online catalogues which are currently based on AACR. Overall, users will probably not notice much difference, although some details might change. RDA has developed an ontology for categorising information content and carrier types, in collaboration with ONIX, the publishing trade organisation (5). This has been used to refine the AACR labels for resource formats to make a clear distinction between content and carrier formats, and this will be reflected in the "physical description" metadata.

A mapping between RDA and the MARC communication format used in most AACR-based online catalogues has revealed only a few mismatches. Full exploitation of the RDA/ONIX framework requires some amendments to the MARC fixed-field codes for content types. This would allow users to restrict searches to general or specific classes of content, for example images, and improve machine-to-machine interoperability between MARC and ONIX-supplied metadata.

RDA would of course have a bigger impact on catalogues based on under-developed metadata content standards. Users would find a significant improvement in the consistency and coherency of the content of the catalogue records, and greater effectiveness in searching for related resources.

The biggest impact that RDA is likely to have on online catalogues is secondary. RDA is closely modelled on FRBR; it uses FRBR terminology where possible, it is structured to match the order of FRBR entities, attributes and sub-attributes, and RDA elements are related to the FRBR user tasks of finding, identifying, selecting and obtaining intellectual and artistic works, realisations of those works (expressions), embodiments of those realisations (manifestations), and exemplars of those embodiments (items), as appropriate. RDA may stimulate the "FRBRization" of

catalogues, and is designed to improve the benefits of catalogues based on the FRBR model.

The FRBR model brings significant benefits for the end-user of an online catalogue. The FRBR entities for the products of intellectual or artistic activity are Work, Expression, Manifestation and Item, as defined above. These entities are related in a mono-hierarchy, from Work to Expression to Manifestation to Item, which can be exploited to provide a hierarchical navigation environment for browsing search results. For example, the results of a title search can be Work-level metadata which can then be expanded to include specific Expressions, then Manifestations, and finally Items. This approach also removes the large amount of duplicate metadata that is often displayed after a search, leading to cluttered displays which are difficult to interpret. For example, many large library catalogues currently display multiple brief records with the same title when the collection contains multiple expressions and manifestations of the same work, leading in some circumstances to search results consisting of pages of indistinguishable metadata.

However, such benefits can be impaired if the metadata content is of insufficient consistency and completeness, as shown by the OCLC FictionFinder prototype (6), so RDA has an important role to play in ensuring the full impact of FRBR on users.

Wide-scale use of RDA will have a significant impact on union catalogues and other metadata aggregations. Physical aggregations created by sharing cataloguing services, copying local metadata to a central catalogue, or harvesting metadata from local repositories will benefit most from more consistent and coherent metadata. Distributed union catalogues would need to develop common approaches to the indexing of metadata content to retain such benefit. Previous research has identified significant variation in the local mapping of metadata fields to specific indexes, such as the inclusion of alternate, part, parallel and related titles in a title index (7). RDA should be able to inform and encourage the adoption of community-wide indexing policies because of its guidelines on the use of descriptive elements as access points and rules for the content of authority files.

Conclusion

Although RDA is unlikely to have any direct impact on library metadata systems and services in the immediate future, it is being designed and developed to have the widest possible take-up at international and national level. It is a significant development in the globalisation of information retrieval services, fitting well within a suite of recent and emerging international standards, and it is that context that its impact will be felt over the next five years.

References

(1) Resource description and access. Available at:
<http://www.collectionscanada.ca/jsc/rda.html>

- (2) Functional requirements for bibliographic records. Available at:
<http://www.ifla.org/VII/s13/frbr/frbr.pdf>
- (3) Managing digital assets in tertiary education toolkit. Available at:
<http://mandate.cdlr.strath.ac.uk/>
- (4) Resource description and access: prototype demonstration. Available at:
<http://www.rdaonline.org/>
- (5) Distinguishing content from carrier: the RDA/ONIX framework for resource categorization. Available at:
<http://www.dlib.org/dlib/january07/dunsire/01dunsire.html>
- (6) OCLC FictionFinder. Available at: <http://fictionfinder.oclc.org/>
- (7) CAIRNS Project recommendations for a cataloguing and indexing strategy for Scottish libraries. Available at: <http://cairns.lib.gla.ac.uk/docs/CAIRNSCatStrat.pdf>
- (8) Improving interoperability in distributed and physical union catalogues through co-ordination of cataloguing and indexing policies: report for Work Package B of the JISC CC-interop project. Available at:
<http://cdlr.strath.ac.uk/pubs/dunsireg/CCICatInterop.pdf>