The future of the MARC: a conversation with Sally H. McCallum

Angela Contessi
Alejandro Gadea Raga

Coordinators GUMARC 21
info.gumarc21@unicatt.it

Abstract

In this interview Sally McCallum, Chief of the Network Development and MARC Standards Office at the Library of Congress, shares with the coordinators of GUMARC 21, the Italian MARC 21 Users Group, her thoughts on important topics related to the future of the MARC 21 format and other emerging standards. Starting from the undeniable widespread implementation of MARC 21 format, Sally McCallum explains the reason why MARCXML was created to provide definitely MARC 21 with flexibility and enable it to work across the "XML divide". She anticipates that MARC 21 will continue to evolve as technologies evolve and that the community's challenge will be to retain continuity with current data. In particular the questions range from the relationships between MARC and formats like MODS, MADS and DC to the possible changes required to MARC 21 to support compatibility with RDA, the new international standard.

Introduction

These days, when the use of markup languages is highly widespread, there are rumors that predict the gradual reduction in use or even the death of the MARC formats. According to these rumors, the Library of Congress, which together with the Library and Archives Canada and the British Library is responsible for the maintenance of the MARC 21 – the most popular format in the library world – will cease to develop this format in 2010.
In fact the present situation looks very different from these predictions. According to the data coming from all over the world, it is clear that MARC 21 is not in decline, but rather in good health. In recent years, the abandonment of national MARC formats and the adoption of MARC 21 as international format has contributed *de facto* to the success of the American format, which is gradually being accepted by an increasing number of users. The reason is probably the fact that MARC 21 belongs in a sense to its users, namely to those who exploit its potentiality.
As a matter of fact, this format implies mechanisms that make possible and even desirable the participation of the library community in its development.
Aware of some recent confusion about the format concerning cataloguing, bibliographic control and bibliographic data exchange, as coordinators of GUMARC 21, the Italian MARC 21 Users Group, we have addressed a few questions to Sally McCallum, Chief of the Network Development and MARC Standards Office at the Library of Congress, to clarify the future of the MARC 21 format. The result is an interesting conversation that we propose below.

1. Last year the Library of Congress, which maintains the MARC 21 formats, published two important updates; now the MARC 21 website offers not only the concise version of the formats but also the full one. In the last five years different European countries like Germany and Spain decided to adopt MARC 21. In front of an undeniable vitality of the

MARC 21 formats, rumors continue to announce the MARC 21's murder. What do you think about this?

MARC 21 is a very mature format, as can be seen by its widespread implementation. This installed user base is what makes it so attractive for continued use and for adoption. However, as we moved into the markup language era, we needed an XML version of MARC for use with newer protocols, and for that reason MARCXML was created. MARCXML takes the MARC 21 data elements, coded values, and data tagging and enables them for XML. Only the MARC structure is left behind so that the beauty of the MARCXML is that is can "play" well with older systems and system modules, via a transformation to the MARC 21 structure, and also with newer XML-based applications. MARCXML is an application of a recently developed ISO standard called MarcXchange (ISO 25577).

The new web publications of MARC 21, the full format and a concise version, serve MARC 21 as well as MARCXML, since for MARCXML it is only the Directory component of MARC 21 that is not needed.

2. With the implementation of RDA: Resource Description and Access, the new international standard (release scheduled for third quarter 2009) will there be changes required to MARC 21 to support compatibility with RDA and ensure effective data exchange into the future? In your opinion, is the success of MARC closely connected to that of the new code?

The international RDA-MARC Committee has been working for the last year on changes needed to MARC 21 in support of characteristics of RDA. The committee is described on the RDA web site and has published discussion papers and proposals on the MARC web site and reviewed them in two Meetings of the MARBI committee. The RDA-MARC Committee members do not think that the proposals will be the end of possible needs, however. They will watch closely the testing of RDA expected to take place the last part of 2009 and early 2010, and then when implementation starts to take place they will be ready to analyze situations to make any additional changes. I expect systems to evolve as they always have, but in this case it will be to make it convenient to use RDA rules, and the communications format will not necessarily need to change. I do not think that we can yet predict the changes that may take place in the exchange environment.

I think that it is necessary for MARC to be supportive of RDA, but we must also recall that MARC must be as rule-neutral as possible -- which it cannot be entirely. There will always be other communities using their own rules with MARC 21.

3. What is the future of the MARC 21 in the USA and in the other countries?

MARC 21 is very fundamental to the data interchange environment in the US and increasingly in other countries. This data exchange has been critical to lowering the costs of cataloging in the US by enabling cooperation on a nationwide scale. It is also critical to the interlibrary loan processes that make one library's resources sharable with another library. American libraries cannot give up those savings -- and especially in a period of shrinking budgets. In addition, American libraries have large investments in automated systems at this point. While MARC 21 is only an exchange format, those systems are highly tuned to the MARC 21 format for data loading and data creation. With shrinking budgets, libraries are more likely to put new interfaces on the MARC-based "back office" catalog than invest in complete new systems. I would think that other countries have some of the same needs and directions, especially since

they have recently been trying to better align their formats with MARC 21 in order to take advantage of system and data exchange opportunities.

MARC 21 is also a suite of formats that drive a variety of library functions, not just descriptive cataloging. The major example is the holdings format that supports, in addition to detailed holdings, serial check in and claiming. Community Information is another format with special applications. In these areas it may be the only format currently available for the function.

That said, I anticipate that MARC 21 will continue to evolve as technologies evolve. The community's challenge will be to retain continuity with current data.

4. In an interview made in Buenos Aires about three years ago, you said that the detail is both the strength and the weakness of MARC 21. Do you still confirm that?

I personally have a "love-hate" relationship with the detail in MARC. I know why it is there. As the community of MARC users increased to take in different cataloging traditions -- archives, cultural materials, non-AACR rule users -- and to describe more and new media with different characteristics -- from sound recordings to maps to moving images to digitized material -- the detail in MARC 21 had to grow to satisfy the description needs for those users. But it made the format "look" very large and has been a deterrent for its use by some groups. If one thinks about it, there is probably a small subset of MARC 21 elements that is used in 70% of the applications. However, MARC 21 detail enables the specialized collections with focused needs to record much more detail than most general applications can afford to do, yet the use of a common format means that the detailed records can be integrated with simpler ones for retrieval and other applications. Perhaps what is needed is to encourage the development of subsets or profiles for applications so that a user group is not confronted with all of the data elements. We have tried to indicate this possibility by putting the "MARC 21 Lite" profile with core elements on the MARC 21 web site.

5. Some librarians believe that new formats like MODS, MADS, etc. can be developed without referring to MARC 21.
    Is this correct?

Formats like MODS and MADS definitely should not be developed without reference to MARC 21. They are intended to be evolutionary from MARC 21, so they accommodate the essential parts of MARC 21 -- the parts that my estimate of 70% of applications use. This MARC 21 compatibility assures that the scholarly community does not lose the ability to find ALL material in a consistent manner, whether the description originated in MARC 21 or MODS. Libraries must provide access to future resources along with the resources from the past.

Libraries have enormous investments in MARC data -- MARC-based systems, MARC 21 records, MARC 21 authority files, and MARC 21 training -- and they need an evolutionary pathway to newer metadata structures. MODS can provide better XML compatibility and different opportunities for use while also assuring that MARC 21 records can be easily transformed into it with little loss.

We, and others in the community, are engaged in modelling with MARC and MODS that could inform the use of semantic web technologies in the future. Retaining relationships from MARC to MODS and then to future formats are important to assure that our data will be able to move forward.

6. Do you think that MARCXML can improve MARC records flexibility and interoperability?

MARCXML definitely provides MARC 21 with flexibility and enables it to work across the "XML divide". One direction that MARCXML could take is to use XML features such as attributes to enhance data in ways the ISO 2709 format structure cannot. These enhanced records might not be totally convertible to the ISO 2709 MARC 21 structure without loss but the loss could be minimized. As noted earlier, MARCXML is also more readily converted to other XML-friendly formats -- internal or local or even RDF/XML. It provides a roadway for the community's enormous storehouse of valuable metadata, the entry points for our collections.

7. Do you think that DC or local formats in XML could be at the same level of MARC 21, and therefore improve the exchange of international bibliographic records, or at the contrary they will just make it impossible?

I do not see DC or local XML formats as viable for international exchange of metadata in the detail that MARC 21 provides. Dublin Core has value as a very core set of elements that are not highly parsed and can be used for a variety of applications, especially those that do not require professional cataloging. Taking DC to a MARC level of detail would lose that flexibility. Sure, DC also does not provide very much specificity but that is what distinguishes it and makes it useful in certain applications. The original aim for DC was as metadata for the web document header that could be supplied by any document creator and a later development was as a common denominator for merging very different data sets, and these are still important uses and goals for DC. As for local XML, we have plenty of experience with that at LC and as a result we also have vast experience mapping data to try to get common data sets for integrated applications without losing consistency or specificity. An international exchange of local XML is not something that we would find useful or cost effective.

Biographical information:

Sally McCallum is presently Chief of the Network Development and MARC Standards Office at the Library of Congress, the Office responsible for the maintenance of the MARC 21 formats and a number of other interoperability-related standards such as an XML version of MARC, the Metadata Object Description Standard (MODS), and the Z39.50 and Search and Retrieve via URL (SRU) Information Retrieval protocols. The Office is also the official home for the Encoded Archival Description XML DTD; the Metadata Encoding and Transmission Standard (METS), and several supportive technical metadata standards; and PREMIS, the emerging standard for preservation metadata for digital objects. She has been an active participant in many organizations and working groups over her more than 20 years at LC, including the MARBI Committee of the American Library Association; boards and committees of the National Information Standards Organization (NISO); and committees of the International Organization for Standardization (ISO) that develop standards for libraries and information services. She has also been very active in the International Federation of Library Associations and Institutions (IFLA), chairing the Professional Board and the Standing Committee on Information Technology and serving on format related committees responsible for the UNIMARC format. She has published a number of articles on standards and networking. McCallum has a BA from Rice University and an MLS from the University of Chicago.