The Deutsche Nationalbibliothek Reports: Germany and Austria’s Migration Project ‘Moving to MARC 21’

VERONIKA LEIBRECHT

Background

The question of the internationalization of library standards has been a point of discussion in Germany for a number of years. Although financial rationalization was always a consideration, the primary motivating factor for the decision to proceed was improved usability for the library user. It is in everyone’s interest that the richness of the bibliographic data developed and gathered worldwide be made accessible to information seekers through familiar, user-friendly and homogeneous systems (Gömpel and Niggemann, 2002). Although the data exchange facet of the international objectives does not directly address the issue of consistent information searches for users, it lays the groundwork by ensuring that bibliographic data is available universally. The migration to the common data exchange format MARC 21, that is improved record sharing between libraries, is a major step towards that goal. The German National Library (Deutsche Nationalbibliothek, DNB), which under legal deposit collects all German and German-language publications from 1913 on, is a service provider first and foremost for German libraries, publishers and the book trade. After the changeover, the many libraries across Germany with special collection mandates will also be able to receive records from abroad through MARC 21 and thus be in a much better position to complete the data in terms of international material offered to users. However, through the internationalization of standards, the Deutsche Nationalbibliothek would thereby also be able to offer its data to virtually all libraries which collect similar material worldwide, offering enrichment to relevant collections internationally.

In December 2004, the Committee for Library Standards (Standardisierungsausschuss)1, which represents library institutions and networks of Germany, Austria and Switzerland (Gömpel, 2009a), confirmed its readiness to improve the exchange and interoperability of data by introducing major changes in national library standards. As a part of this strategy, which would also look at adopting international rules for description and access, it was decided to move from the national data exchange format, MAB2 (Maschinelles...
THE DEUTSCHE NATIONALBIBLIOTHEK REPORTS

Austauschformat für Bibliotheken) to the international standard MARC 21 in Germany and Austria, Switzerland having completed its changeover to MARC 21 in the 1990s. In the introduction of the community-wide application of the standard, the German National Library\(^2\) and its Office for Library Standards would take on the responsibility for the overall project coordination and the Expert Group for Data Formats (Expertengruppe für Datenformate) the development and maintenance of bibliographic data formats.

On 30 June 2009, Germany and Austria’s migration to MARC 21 project came to completion, meaning that the groundwork for the implementation of the format in Germany and Austria was laid and national and international data exchange in the MARC 21 format could begin. The project’s successful completion over the past five years has depended on the close cooperation of many partners. Funding has come from the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG)\(^3\) and the Andrew W. Mellon Foundation, procedural decisions and organization from the Committee for Library Standards and the Office for Library Standards and the Expert Group for Data Formats the MARC Standards Office at the Library of Congress and the Machine-Readable Bibliographic Information committee (MARBI).

PROJECT OUTLINE

The core of the technical preparations for the changeover were, firstly, the mapping of MAB2 to MARC 21 and the design of an editorial support system at the Deutsche Nationalbibliothek affectionately named POLLY (Platform for Operational Layers in the Library), to hold the translation of MARC 21 into German. The third cornerstone of the migration revolved around public relations work. Strategies for supportive communication included the translation of the standard into German, workshops for experts and the general library community and systems developers and cultivation of contacts with international committees and experts.

The complete mapping from MAB2 to MARC 21 was considered to be the central applied technical element of all combined efforts in the project. Through it, gaps between the two formats could be identified and decisions made about where, for bibliographic information deemed absolutely necessary for the German-speaking library world to transport, official changes to MARC should be proposed. Locally-defined XX9/X9X/9XX fields would be conceived or filled for idiosyncratic information transport needs.

POLLY, designed to handle the translation of the MARC 21 standards, can process and edit virtually every type of XML-based document. Essentially, it is a combination of a content management system (for searching and management of text), publishing application (for editing and validating) and browser-based community tool (with annotation features and variable rights settings). Although it was developed with the translation of the text of MARC 21 in mind, it is hoped that it will be useful for further translations such as RDA in the future. Because the Deutsche Nationalbibliothek is its owner, it is possible to grant rights to the application to third parties.\(^4\)

Important to the project in the German and Austrian library community was our commitment to acceptance through regular and extensive communication not only with the experts directly involved in the preparatory work for the

\(^2\) Because of historical circumstances in Germany, a national library was not founded until 1912 with the Deutsche Bücherei in Leipzig. After 1945, an additional location was established in Frankfurt am Main for the western states of Germany and in 1970 the Deutsches Musikarchiv in Berlin became a part of those. On the occasion of the German reunification these institutions were brought together to form ‘Die Deutsche Bibliothek’, which in 2006 obtained both an expanded legal mandate and a new name: Deutsche Nationalbibliothek (German National Library) (Ansorge, 2009).


\(^4\) Further information about the editorial support system POLLY can be found at the ‘Moving to MARC 21’ website.
migration, but to the broader library community. The primary means of information dissemination were the DNB website and its special pages about the project, e-mail lists (the DNB Datenformate list and the Library of Congress MARC Forum), training and information events in the form of workshops and symposia for experts (described below), presentations at conferences such as the Deutscher Bibliothekartag and articles in major German journals and newsletters such as Dialog mit Bibliotheken, Zeitschrift für Bibliothekswesen und Bibliografie and Newsletter Standardisierung und Erschließung. We aimed to inform all interested or affected parties both nationally and internationally to their satisfaction.

Above, I have listed the codes and briefly described the decision-making processes for the German and Austrian system of libraries. As for an introductory description of additional aspects of the German and Austrian library landscape, especially those concerning state-level organization, for the purposes of this article, only the complexity of the German language, as described in Mark Twain’s essay, The Awful German Language, can rival its convolutions. Twain said that his philological studies satisfied him that a gifted person ought to learn English (barring spelling and pronouncing) in thirty hours, French in thirty days, and German in thirty years (Twain, 2004). Because of the federal system in Germany, consortia include and represent libraries which existed long before the German state as such and therefore have traditionally and independently held the written cultural heritage of their individual counties, duchies, kingdoms and so on (Hillmann, 2009). Without background knowledge, the task of learning about the hierarchy and web of organizations, networks, associations and consortia, which often, due to their historical development, have highly structured membership arrangements with members located at opposite ends of Germany can seem perplexing to those outside of Germany and may initially seem as daunting as the task of acquisition of the German language with its lengthy composite nouns and inflected grammar. I will go under the assumption that succinct translations into English of the function of organizations will suffice and therefore keep explanations simple for the purposes of this article.

THE GERMAN AND AUSTRIAN LIBRARY LANDSCAPE

For the purposes of following the information in this article, it may be useful to know the names of the rules and formats which are in use in Germany. For bibliographic description, the main cataloguing tools are the rules for descriptive cataloguing (Regeln für die alphabetische Katalogisierung, RAK) and for subject indexing, the rules for subject headings (Regeln für den Schlagwortkatalog, RSWK). Exchange procedures are executed by means of the machine-readable data exchange format MAB, developed in 1973 (Altenhöner, 2009). The Expert Group for Data Formats has been responsible for the development and the maintenance of the MAB format under the auspices of the German National Library. The latest format version MAB2 is a national data exchange format and therefore restricted to Germany and Austria. Due to the intensity of the work involved in the migration and the priority given to MARC 21 for future data exchange, there have been no changes in MAB since the resolution of December 2004 to adopt international standards – MAB was stabilized. It is in German and Austrian libraries’ best interest to become fully operational in MARC as soon as possible for a current and complete transport of data.

The preparatory work for the migration began in 2005. On the basis of a comparison of both data exchange formats, the Expert Group for Data Formats devoted several meetings working out
THE DEUTSCHE NATIONALBIBLIOTHEK REPORTS

critical points relevant to the mapping process. Early in 2006, the mapping process between MAB2 and MARC 21 was set in motion. It shows on the one hand where MAB2 data appears in MARC 21, and on the other hand where the mapping fails.5 Dealing with the resulting gaps and deficiencies was one of the main tasks of the German experts. The main focus lay on three aspects: the mapping of hierarchical structures in the context of multivolume works (multipart monographs, series), the connection between headings and appropriate authority files, and sorting rules for volumes. All aspects are part of the German library tradition and crucial when it comes to data exchange; they therefore raise the question, if and how it is possible to transport the connected information without a loss. Keeping in mind that the stability and consistency of MARC 21 were a high priority, the experts got more acquainted with the MARC format and in many cases were finally able to find solutions either by including the needed format elements into MARC 21 or at last by defining them on a national basis. By doing so, the experts minimized the amount of extensions required in order not to stress the extensibility of the format. Only the remaining data elements considered absolutely necessary for the information transfer were collected in discussion papers and presented at the meeting of the MARBI Committee. At the MARBI committee, Germany and Austria were successful with most of the requested changes and consequently they were built into the concordance and became the cornerstone for the changeover and processing routines in the libraries.

As it turned out, a large part of the German data could be manifested and transported in MARC 21. Although there was plenty of support from experts such as Sally McCallum at the Library of Congress, a certain amount of ‘home study’ necessarily took place. An interesting by-product of this is that MARC is used in a strictly compliant way, but MARC conventions are not always followed. Running a small sample of data through software called ‘MARC Report’ at the Germany National Library, we got feedback on the ‘MARCness’ of our records. For example, although it is perfectly legitimate to repeat field 240 in MARC Bibliographic, the repeatable use is an example of what up until recently has not been considered common practice at OCLC and the LC and thus attracted the attention of the programme. This experience resonates with findings of studies which have found that in practice a limited range of available MARC content designation structures are in fact used (Moen and Miksa, 2005). When sending out Personal Name Authority data to the Virtual International Authority File (VIAF), for example, we included a document describing which fields in MARC we use and how, in order to assist those interpreting the records. These can be found for both bibliographic and authority data on our website (Gömpel, 2009b). In this case OCLC, in processing our data, will need to make some adjustments in order to load our records into World Cat, as compatibility is near perfect but naturally not absolute.

The proposals to MARBI were divided into three main parts. The first part dealt with those concerning MARC Bibliographic and contained three blocks: one block for a new subfield for record control numbers which are designated for linking purposes in the context of multivolume works and authority files, one block for the use of sorting rules for volumes, and one block for diverse aspects, for example codes, numbers, and notes. The second part contains one proposal concerning MARC Holdings, and the third part deals with proposals belonging to MARC Authority. Here the focus is on the implementation of a subfield for record control numbers and the implementation and reshaping of fields originally belonging to MARC Bibliographic in MARC Authority which should be used for language codes or dates/time periods. Half a year later, proposals for the representation of Dewey Decimal System data followed (Heuvelmann, 2007).

The German authority files PND (Personal

5 The mapping results can be accessed at the ‘Moving to MARC 21’ website.

6 http://www.marcofquality.com/soft/softindex.html
Name Authority File), GKD (Corporate Body Authority File) and SWD (Subject Heading File) are held by the German National Library. They are updated, edited and managed in cooperation with the DNB and the German regional library networks. The networks provide catalogue databases on the regional level in which mirrors of the national authority files are integrated. ‘Integrated’ means that the authority control numbers are included either in the appropriate access point field of the bibliographic record or in the appropriate index for the access point (via the control numbers search queries can be filed simultaneously to the authority files and the bibliographic files). The DNB provides a weekly update service for the national authority files which the regional networks and many single users import as routine updates. In these update processes the authority control numbers serve as routing elements, not only for the authority records but also for the bibliographic records.

The extensiveness of the German authority files and the consistency of linkage to relevant entities in bibliographic records are a particularity of the German-speaking world.7 Presently, the three authority files are being merged and MARC 21 is proving to be well suited to carry its information structures. With the changes made to MARC 21 to accommodate this system, the blueprint for linked data was elaborated, which on the one hand speaks for the flexibility of the format to handle and reflect conceptions of the future. On the other hand, it begs the question of the viability of its longevity in an information environment where semantic web models link data through Uniform Resource Identifiers (URIs) with the Resource Description Framework (RDF) serving as the connecting structural code.

INFORMATION EVENTS

Because the need for broader and deeper knowledge in MARC 21 for the library community was recognized, a total of five information events were offered at the German National Library during the course of the project (Gömpel, 2009b)8. The first workshop was held from 11 to 12 July 2005 for the German and Austrian library community. Randall Barry from the Library of Congress in Washington agreed to lead the workshop. A full day of introductory presentations to MARC 21 was followed by a second day on which participants were given the chance to delve into the details of implementation and discuss issues concerning the format in a smaller circle of experts. Such topics as basic materials cataloguing in the bibliographic, authority, community and holdings formats, content designation, new technologies and mapping to MAB were presented and discussed.

For the benefit of those directly involved in the implementation of MARC 21 in their libraries, namely the members of the Expert Group for Data Formats, Sally McCallum, Chief of the Network Development and MARC Standards Office at the Library of Congress, was invited to lead discussions from 27 to 29 September 2006 at the Deutsche Nationalbibliothek. The discussions around the mapping of MAB to MARC 21 became the basis for the analysis of the bibliographic elements traditionally transported in MAB, but which could not be adequately depicted in MARC.

An indispensable aspect of the move to MARC 21 was the sharing of experiences by other national libraries made in the process of their migration. This was discussed at an information event named ‘Experiences, challenges and visions...’ where Anders Cato, Mirna Willer and Alan Danskin spoke about the move for Sweden, Croatia and the UK respectively.

On 26 September 2007, vendors of library systems were invited to the DNB along with the Expert Group for Data Formats to be informed of both the latest developments in the conversion to MARC and to share their progress in adapting their software to the future needs of German and Austrian data transfer.

7 The PND has 3.4 million records, 1.7 million of which are personalized. The GKD and SWD have 1.2 million and 560,000 records respectively.
8 Powerpoints of the presentations can be accessed at the ‘Moving to MARC 21’ website.
Having covered the introduction of the application of MARC for librarians and data format experts as well as vendors of library systems, the project and series of workshops came to a close with the symposium ‘MARC–ing a new landscape’ on 2 June, 2009. Presentations were given by Gabriele Meßmer of the Bavarian Library Network (Bibliotheksverbund Bayern, BVB), Cornelia Katz of the Library Service Centre of Baden-Württemberg (Bibliotheksservicezentrum Baden-Württemberg, BSZ) and Reinhold Heuvelmann of the German National Library, all directly involved in the project. Three recognized international MARC experts were also invited to speak: Sally McCallum of the Library of Congress, Margaret Stewart, at the time chair of the Joint Steering Committee for Development of RDA, a member of the Canadian Committee on MARC, the MARC Advisory Committee and MARBI, and thus an expert on data format and cataloguing rules, and Tom Delsey. Tom Delsey has held senior positions in the Canadian National Library for more than twenty years, and is the Editor of Resource Description and Access.

Reactions to the above-mentioned symposium from other national libraries presently embarking on a changeover to MARC indicated that the event was helpful for their own project not only as a touchstone for the practicalities of such a move, but also as a source of encouragement; the changeover is, after all, a long process. From the perspective of the Office for Library Standards, the symposium also played the role of marking the occasion of the completion of the project. Our thanks went to those who accompanied the process. As well as summarizing what was done to initiate and carry out the changeover, the discussion about the past raised questions about the future. Offering food for thought and some guidance around the questions with which libraries are grappling currently, the questions about the future of MARC turned to its suitability to reflect and transport the principles of RDA. As with MARC 21 in 2004, the German Committee for Library Standards will be considering the benefits and appropriateness of phasing out RAK and RSWK and adopting RDA once it is available in its full online glory. At the symposium there was an opportunity to learn about the many ways in which MARC 21 can reflect the principles of RDA.

Communication about a project in a country where the language is not English adds an extra dimension to organizational processes because it entails informing professionals and providing tools in at least two languages to cover both the national and the international communities, as well as taking cultural differences into consideration (described in this article primarily in terms of conventions in library science). In the case of the transition to MARC 21, librarians in Germany and Austria received information on the subject matter in their native tongue, whereas communication with our partners at the Library of Congress and MARBI was carried out in English. From the Office for Library Standards to the Expert Group on Data Formats, sources of information related to the migration, both in terms of documents and contacts, needed to be understood, analysed and responded to in English. This is true for everything from the MARBI discussion papers and proposals to presentations at conferences and workshops such as the American Library Association conference or Germany’s largest annual library conference, Deutscher Bibliothekartag. For the most part, experts in data transfer, whose professional profile amongst other international work includes networking, the challenge, which requires flexibility and certainly some mettle, seems to have been welcome. However, it was felt that librarians who would be dealing with data transfer on a routine basis and are accustomed to working in German should continue to be able to do so. For this reason, a translation of MARC 21 into the working language of those who would be using or potentially using the standard was commissioned. Even if many librarians have adequate English in Germany and Austria, the translation contributed to an acceptance of the project.

**TRANSLATION ISSUES**

As we are also seeing at the moment at the Office for Library Standards in our translation of
THE DEUTSCHE NATIONALBIBLIOTHEK REPORTS

elements of RDA, the experience with MARC 21 proved that standardization of target-language vocabulary is a primary condition for translation, to be determined by teams of librarians from various areas of speciality assigning subject headings, descriptive cataloguing and Information Technology (IT). Although it is time consuming, a process of reaching consensus through experts ensures best reliability and long-term acceptance by the community. Translators specializing in library science are few and far between and even so, could not be expected to appropriate and apply the vast vocabulary in the field, especially as regards the steady flow of neologisms into the pool of terms over the years. The challenge lies in matching the text to be translated with already existing terms in the target language, while allowing for some flexibility for the formation of new terms. Regarding new expressions in translation, their conception is not so very much different from writing original texts about new subject matter within one language, as one can see from the coinage of such terms as expression to expression relationships and content type in Functional Requirements for Bibliographic Records and RDA respectively. The difference being that when we translate, the concepts are a given in the source language, even if determining their connotation often can still be cause for head scratching. As with the translation of a classification, the language and culture is forced to fit the conceptual structure of the dominant language (Hudon, 1997). Ideally, for instance in the case of the Moving to MARC 21 project, one should be trained in both Anglo-American and German library science while speaking both languages fluently. In German, such hard to come by versatility is embodied in the term ‘eierlegende Wollmilchsau’ or an egg-laying wool and milk sow (Wikipedia, 2009).

The Deutsche Nationalbibliothek assigns Dewey Decimal Classification Numbers to its collection (for the purpose of information searches but not collocation) and uses the full Dewey Decimal Classification which has been translated into German in both print and online editions. The translation of MARC fields relating to the DDC in the bibliographic format needs to be in line with the official translation of the classification. In the case of updates for MARC 21, the most recent release being number 9 of October 2008, MARC-specific terms for the fields 083 and 085 such as ‘internal subarrangement’ and ‘root number’ recently had to be coined by the German editors of the DDC at the Cologne University of Applied Sciences (Fachhochschule Köln) since they do not exist in the official translation of the DDC. This will continue to be a challenge as updates are released.

The target language of MARC translations at the DNB is not always German. In the case of the locally defined fields, definitions of used fields need to go to those importing our data from abroad. In those cases, we are translating what we do in Germany into English. Questions arise which can require extensive research and analysis and lead to discussion amongst German librarians. Is a Schlagwortkette in the RSWK the same thing as a subject heading string for Library of Congress Subject Headings if they are formed by different rules? Do we need both the term ‘occupation’ and ‘profession’ to cover the meaning of ‘Berufsbezeichnung’ in the German Name Authority File? And, will the matter at hand be clear to our Anglo-American partners?

LIBRARY DATA EXCHANGE TRADITIONS IN AUSTRIA AND GERMANY

Not only language, but the cultural heritage of library standards came into play and added to the complexity of the issues at hand; although there were common denominators with Anglo-American language and cataloguing traditions, they needed to be analysed and identified, while the irreconcilable differences needed to be labelled and addressed. It should be noted here that a major procedural difference between the general modus operandi of the Anglo-American and the German or Austrian cataloguer of most of the networks is that for the latter, the data exchange format is irrelevant to daily routines. Austrian and German cataloguers would not be speaking the same cataloguing language even if the conversation took place in English and both had a list of fields from MARC and MAB mapped to each
other. Many German and Austrian cataloguers do not enter information according to MAB tags. The advantage of this is that tags specific to the software of choice of the given library or library network have been designed especially to facilitate the cataloguing process through user-friendly input options and few subfields. From the point of view of IT, there is more flexibility for the manipulation of the data in the specific cataloguing software, whereby the currency of data exchange format fields is less of an issue at the point in time of data entry. Records made by the cataloguers are automatically converted into an expanded version with subfields, which is in turn converted to MAB behind the scenes only when it is for export, or vice versa, import of data. Essentially, this means that the implementation of MARC was handled in Germany to a great extent by librarians who work specifically in data processing and who are responsible for completing the behind the scenes preparation of mapping MAB to MARC 21 and the handling of interfaces. The consciousness of this difference in the cataloguer’s tools has been a learning process for both the Germans/Austrians and Anglo-Americans and requires flexibility in terms of understanding cataloguing issues on both sides of the ocean when discussing MARC. Having said that, some few library networks presently catalogue in MAB and inspired by the migration project’s mandate for a more international outlook, will be implementing a workflow where cataloguing takes place directly in MARC in the near future.

CONCLUSION

The Deutsche Nationalbibliothek is currently receiving data from the British Library and the Library of Congress and delivering to OCLC for World Cat and VIAF (Virtual Name Authority File) and to The European Library (TEL) database in MARC 21. Within Germany, two networks are already taking data from the DNB in MARC 21, while others plan to be at that stage in the near future. Although we previously exchanged data internationally, the project is an important component of continued and future use of international standards. Since more than one format had to be operated in the past, the migration will serve to reduce the effort needed for conversion work. While costs are reduced, international compatibility has been established and data can now be more efficiently and widely processed.

As was said in the project-closing symposium, we are indebted to those who supported us on our way. But it is also our hope that the international cooperation has been beneficial to partners. The enrichment of the format through the elements added through MARBI can raise awareness of possibilities which did not exist before or could be useful to others in such concrete areas as record linking for authority data in bibliographic records. The changeover has also contributed to the library community by offering all partners increased exchange of comprehensive records in their respective areas of specialty and the benefit of copy cataloguing. The scope of the Moving to MARC 21 project, although designed for the concrete goals of increased international bibliographic data exchange and thereby ultimately the improvement of the quantity and quality of information for the user, which from the German and Austrian perspective at least it most certainly has accomplished, has also served to support the broader interest of an international exchange of language, bibliographic practice and ideas. Was it worth it? We say ‘yes’ and ‘ja’!

REFERENCES


THE DEUTSCHE NATIONALBIBLIOTHEK REPORTS


ABSTRACT

With the aim of improving international bibliographic interoperability, in 2004 Germany and Austria initiated a migration from the German data exchange format Maschinelles Austauschformat für Bibliotheken (MAB2) to the Anglo-American format MARC 21. The article outlines the main elements of the project ‘Moving to MARC 21’: a mapping of the data exchange formats MAB2 and MARC 21 to one another to ascertain gaps between the formats, the design of an editorial support system for the translation of MARC 21, and professional information sessions introducing the new format to the library community. Elements from MAB2 considered to be indispensable for exchange were either introduced to the format through Machine-Readable Bibliographic Information committee (MARBI) proposals or where they would not need Germany-wide application, defined in local fields. The project came to completion in June 2009 and has laid the foundation for international data exchange by assimilating processes and eliminating the need for intensive conversion work in Germany and Austria. An incidental but significant aspect of the project was the productive combined efforts of German and Austrian professionals and their Anglo-American partners in successfully overcoming differences in language and bibliographic transport and processing traditions through cooperative problem-solving.

ALEXANDRIA, 21(1), 2009