

Arbeitsstelle Datenformate

Status: verabschiedet

Protokoll der 14. Sitzung der Expertengruppe Datenformate am Mittwoch, 3. Dezember 2008

Deutsche Nationalbibliothek Frankfurt, Adickesallee 1, Tagungsraum

Teilnehmer

Reinhard Altenhöner (Vorsitz)	Deutsche Nationalbibliothek
Dr. Barbara Block	Verbundzentrale des GBV
Reinhold Heuvelmann	Deutsche Nationalbibliothek
Angelika Hinrichs	Hessisches Bibliotheks-Informationssystem
Peter Kassner	ekz Bibliotheksservice GmbH Reutlingen
Cornelia Katz	Bibliotheksservice-Zentrum Baden-Württemberg
Josef Labner	Österreichischer Bibliothekenverbund
Gabriele Meßmer	Bayerische Staatsbibliothek München / Bibliotheksverbund Bayern
Beate Rusch	Kooperativer Bibliotheksverbund Berlin-Brandenburg
Karin Schmidgall	Literaturarchiv Marbach, für: Arbeitsgemeinschaft der Spezialbibliotheken
Stephani Scholz	Hochschulbibliothekszentrum des Landes Nordrhein-Westfalen
Barbara Sigris	Staatsbibliothek zu Berlin, für Zeitschriftendatenbank

Entschuldigt

Adrian Nolte	Stadtbibliothek Essen, für: Öffentliche Bibliotheken
Sabine Thänert	Deutsches Archäologisches Institut

Gäste

Christine Frodl (zu TOP 5a)	Deutsche Nationalbibliothek
Christel Hengel-Dittrich	Deutsche Nationalbibliothek
Susanne Oehlschläger	Deutsche Nationalbibliothek

Protokoll

Veronika Leibrecht	Deutsche Nationalbibliothek
--------------------	-----------------------------

Tagesordnung:

1. Begrüßung, Personelles
2. Genehmigung der Tagesordnung
3. Verabschiedung des Protokolls der 14. Sitzung am 16. September 2008.
4. Bericht aus der Arbeitsgruppe Kooperative Neukatalogisierung
5. Umstieg auf MARC 21
 - a. RDA
 - b. RDA/MARC Working Group, MARBI
 - c. Mehrbändig begrenzte Werke
 - d. Abbildung von RSWK-Daten in MAB2 und MARC 21
 - e. Abbildung von DDC-Daten in MAB2 und MARC 21
 - f. MARC-Antrag 2008-02: Abbildung der Nationallizenzen
 - g. ISSN-L in MAB2 und MARC 21
 - h. IDN-Transport in MARC 21
 - i. Dokumentation der Anwenderebene
 - j. Weitere Planungen
6. Verschiedenes

1. Begrüßung, Personelles

Der Vorsitzende Herr Alenhöner eröffnet die Sitzung und begrüßt die Mitglieder der Expertengruppe Datenformate und die Gäste. Er entschuldigt Herrn Nolte und Frau Thänert, die an der Sitzungsteilnahme verhindert sind. Zum 1. November 2008 wurde Herr Christian Aliverti als Nachfolger für Frau Uebelhard-Petzold benannt. Er nimmt an der nächsten Sitzung teil.

2. Genehmigung der Tagesordnung

Zum Tagesordnungspunkt 2 liegt den Teilnehmern das Dokument T_DF_20081203.doc "*Tagesordnung*" vor.

Die Tagesordnung wird mit Ergänzungen zu TOP 5, "MARC-Änderungsantrag zu autorisierten ISSN-L" und "Publikation MARC 21" in der vorliegenden Form angenommen.

3. Verabschiedung des Protokolls der 14. Sitzung am 16. September 2008

Zum Tagesordnungspunkt 3 liegt den Teilnehmern das Dokument P_DF_20080312_E2 "*Protokoll der 12. Sitzung der Expertengruppe Datenformate*" in der überarbeiteten Entwurfsfassung vor. Die Korrekturvorschläge von Frau Schmidgall (19. November 2008) werden besprochen. Der vorliegende Entwurf wird außerdem unter 5.c und 5.e geändert und geht mit der Bitte um Rückmeldung bis Freitag den 12.12.2008 in veränderter Form an alle Sitzungsteilnehmer.

4. Bericht aus der Arbeitsgruppe Kooperative Neukatalogisierung

Zum Tagesordnungspunkt 4 liegt den Teilnehmern das Dokument "*Bericht 2. Halbjahr 2008*" der Arbeitsgruppe "Kooperative Neukatalogisierung" der Arbeitsgemeinschaft der Verbundsysteme vor.

Frau Block berichtet, dass das Papier zu MARC 21 "*MARC 21 Spezifikationen: Vereinbarungen der Arbeitsgruppe Kooperative Verbundanwendungen der AG der Verbundsysteme zum Datenaustausch in MARC 21*" zurzeit in der Abstimmungsphase ist. Anfang 2009 wird ein Wiki eingerichtet, das auch den Mitgliedern der EG Datenformate geöffnet wird. 2009 werden Tests mit vorhandenen Daten gestartet.

5. Umstieg auf MARC 21

a. RDA

Frau Frodl stellt sich als Leiterin der Arbeitsstelle für Regelwerke in der Arbeitsstelle für Standardisierung an der DNB vor. Sie berichtet, dass der aktuelle RDA-Gesamtentwurf am 17.11. im PDF-Format in 40 Dateien mit vielen Begleitpapieren vom Joint Steering Committee veröffentlicht wurde.

Das Online-Tool war nicht rechtzeitig fertig gestellt, soll aber im Dezember erscheinen. Für das 3. Quartal 2009 ist die endgültige Fassung angekündigt. Für das Online-Tool wurde ein Testfahrplan entworfen, der z.B. auch für Nicht-Bibliothekare eingesetzt werden kann. Frau Frodl informiert die Expertengruppe sobald weitere Erkenntnisse vorliegen.

Als Bestandteile der Öffentlichkeitsarbeit werden auf dem Bibliothekartag in Erfurt Veranstaltungen zum Thema RDA abgehalten. Außerdem hat die DNB schon einige Übersetzungen ("*cover letter*", zur Veröffentlichung des Full Draft, Vokabular, Prospektus und FAQ) vorgenommen, die auf der Website der DNB veröffentlicht werden. In Planung ist weiterhin die Übersetzung des Glossars.

Die DNB ist seit 2006 an den Stellungnahmen des RDA-Entwurfs beteiligt. Die jetzige Frist für die Abgabe der Stellungnahmen zum Regelwerk an das JSC ist der 2. Februar 2009. Ein RDA-Wiki der DNB beinhaltet Stellungnahmen von DNB-internen Sitzungen zum Regelwerk, sowie Begleitpapiere zur weiteren Information. Die Kommentierungsfrist für das Wiki läuft bis zum 8. Januar. Am 16. Januar kommen die Mitglieder der Expertengruppen zusammen, um offene Punkte zu besprechen, bevor die endgültige Fassung für das JSC formuliert wird. Eine Einladung wird verschickt und gemäß der wichtigsten Wiki-Einträge werden offene Fragen für die Tagesordnung zusammengestellt. Punkte, die noch nicht eingeflossen sind, werden auch im Wiki stehen. Bisher noch nicht beantwortete Fragen und Anregungen der letzten Stellungnahmen sollen laut JSC nochmals geschickt werden.

Frau Block bittet für ihre Kollegen des GBV darum, dass noch diese Woche eine Fassung der Stellungnahme verfügbar gemacht wird, wenn auch unvollständig.

Der BVB trifft sich zur Besprechung der RDA und eine Besprechung mit Experten ist festgelegt.

Für die Kollegen der Verbundarbeitsgruppen wird um Gastleseerkennungen gebeten. Schreibrechte sollten dem Kreis der Expertengruppen vorbehalten sein. Pro Verbund sollte eine Stellungnahme verfasst werden.

Einzelne Teile von RDA sind noch nicht aufeinander abgestimmt (Beispiele, Verweisungen, Definitionen, Zählungen) und bedürfen einer Überarbeitung. Der Anhang M enttäuscht, weil sich keine große Veränderung zur vorhergehenden Version gezeigt hat. Es wird darauf hingewiesen, dass es sich um einen Entwurf handelt.

b. RDA/MARC Working Group, MARBI

Zum Tagesordnungspunkt 5.b. liegen den Teilnehmern die Papiere "*RDA Full draft, chapter 3.2 'Media Type'*", "*RDA Full draft, chapter 3.3 'Carrier Type'*" und "*RDA Full draft, chapter 6.10 'Content Type'*" vor.

Reinhold Heuvelmann betont die Wichtigkeit der Expertengruppe Datenformate zur Unterstützung seiner Tätigkeit als Mitglied der MARC/RDA Working Group. Die EG Datenformate sollte überlegen welche Rolle sie in der Analyse und Auswertung des Regelwerks spielt. Zur Klärung dieser Thematik stehen zwei Modelle zur Verfügung. Das Modell eines DNB-FRBR Workshops (Juli 2004) zeigte die Beziehung zwischen FRBR, *cataloguing code* und *exchange format* noch als Dreieck. Jetzt sind FRBR und FRAD als grundlegende Modelle in die RDA integriert. Die Rolle der EG Datenformate ist es, die Beziehung Datenformat Regelwerk zu behandeln und zu ermitteln, wie RDA in MARC 21 abzubilden ist. Das Dokument "*RDA database implementation scenarios*" (Januar 2007) erläutert drei mögliche Modelle für RDA-Daten. Szenario 3 wird "flach" genannt, weil es keine Links zwischen Datensätzen ist. Dies spiegelt den jetzigen Stand eines großen Teils der angloamerikanischen Datenbankstrukturen wider. Szenario 2 bildet den Ist-Zustand von Deutschland und Österreich ab, betont die Beziehungen zwischen Datenbeständen (z.B. zwischen Titeldaten und Normdaten). Szenario 1 hingegen zeigt einen Blick in die Zukunft in Form einer relationalen, objekt-orientierten Datenbankstruktur. Nach FRBR gegliedert, sieht es *work* als eigene Entität vor, und *expression* als optionale Entität vor. Interessant für die EG Datenformate ist die Frage, wie man das in MARC 21 abbilden würde und wie Felder verwendet werden können.

Zur weiteren Orientierung möglicher Vorgehensweisen in der Abbildung von Beziehungen zwischen FRBR und FRAD Entitäten beschreibt Herr Heuvelmann die zwei vorgeschlagenen Modelle des Papiers von 2005 "*Using MARC 21 with FRBR: Record Configurations*". Sally McCallum hat darin vorgestellt, wie mit *work* und *expression* im Datenformat MARC 21 umgegangen werden könnte. Das im Papier enthaltene "Modell A" setzt Normdatensätze (*Authority*) für *work/expression*-Datensätze ein. Das Gegenmodell B hat zum Teil gegenläufige Tendenzen, weil es *work/expression* nicht in Normdatensätzen sondern in Titeldatensätzen abbildet. Es gibt eine Tendenz in der internationalen Diskussion, *work/expression* eher in *Authority* als in *Bibliographic* abzulegen.

Herr Heuvelmann tendiert zum Modell B, da *work* und *expression* einen bibliografischen Charakter haben.

In dem Papier wird der Unterschied zwischen *internal record configurations*, das eine gewisse Flexibilität in der Definition der Entitäten unterschiedlicher Bedürfnisse von Bibliotheken erlaubt und Austauschdatensätzen, die Entscheidungen zur Struktur für die Gewährleistung der Austauschbarkeit treffen müssen, dargestellt. MARC muss Datensätze mit Verknüpfungen verbinden und gleichzeitig als Datenformat unabhängig von Bibliothekssystemen austauschen können. Da die Grundsätze der Verknüpfung zwischen Datensätzen in MAB2 möglich sind, wird postuliert, dass diese Anforderungen auch in MARC erfüllbar sein sollten.

Die Frage wird gestellt wie viele Datenelemente von *work/expression* im *manifestation* enthalten sind und wie die Entitätenebene im Datensatz gekennzeichnet wird. Es wird angemerkt, dass durch eigene Datensätze für *work* and *expression* auf den Datensatz, wie es in MAB zwischen Titeldaten mit Normdaten gemacht wird, die Identität des Satzes nach FRBR gekennzeichnet werden kann.

Die Fußnoten-Felder in MARC könnten in der Abbildung von RDA eine neue Bedeutung zur (Darstellung) von Informationen aus RDA bekommen. Doch ist das Verschieben von spezifischen Datenelementen in allgemeine Fußnoten nicht erstrebenswert. Spezifische Felder sollten dafür ausgesucht werden. Dieses trifft vor allem zu, wenn man Funktionen zuordnen will und statt *description*- auch *access*-Informationen abgebildet werden sollen.

Die Services, die für den User auf die Daten aufgesetzt werden, sind ein interessanter Aspekt der Datenstruktur. Grundsatzfragen sind u. a. ob MARC in RDA angewendet werden kann, welche Alternativen es gibt und wie die Services, die auf den Daten aufgebaut werden, ausschauen.

Andere Techniken, um RDA-Datenelemente abzubilden sind prinzipiell denkbar: RDA kann auch in Dublin Core oder ONIX dargestellt werden. Das Vorhaben der RDA/MARC Working Group ist es jedoch primär Kompatibilität zwischen RDA und MARC herzustellen. Ein Arbeitspaket der RDA/MARC Working Group ist die Codierung von *content*, *media* und *carrier type*, unter Verwendung des Modells "RDA/ONIX Framework for Resource Categorization" des JSC. Die Vorteile wären eine internationale Austauschbarkeit durch Zusatz von auflösbaren Codes in verbale Bezeichnungen in jeweils einem speziell für die drei Angaben angelegten MARC-Feld.

Theoretisch denkbar wäre es, ein umfassendes und detailliertes Mapping vorzunehmen, um festzustellen, was von RDA und MARC zusammenpasst. Das in RDA enthaltene Mapping von MARC 21 nach RDA hat Lücken, da es nicht auf dem letzten Stand von MARC basiert und nur variable Felder und Unterfelder behandelt. Eine Konkordanz von RDA nach MARC ist nicht Teil des RDA-Entwurfs, soll jedoch noch für das 2. Quartal 2009 erstellt werden.

Die Felder 007 oder 008 sind nicht im Gespräch für Veränderungen, um RDA Elemente abzubilden, da sie kompliziert und aufwändig sind. Der Gebrauch von neuen Feldern wird vorgezogen.

Zur Untersuchung der Eignung von MARC für den Datenaustausch von RDA-Daten sowie die Wahl einer der besprochenen Modelle A und B zur Abbildung der Beziehung zwischen Entitäten wird vorgeschlagen, dass Tests gemacht werden. Indem die Kennzeichnung und Verknüpfung in MARC ausprobiert wird, könnte die Frage, ob es geeigneter ist, *work* und *expression* in Bibliographic oder in Authority zu verankern, besser entschieden werden. Es wird vorgeschlagen eine Datenbank aufzubauen, um Modelle mit einzelnen Sätzen durchzuspielen. Frau Schmidgall berichtet, dass das Literaturarchiv in Marbach 2009 nach der erfolgreichen Migration Daten hat, die in diesem Zusammenhang besonders interessant sein könnten.

Dieser Tagesordnungspunkt wird zusammengefasst mit der Auflistung der Punkte, die von der RMWG bereits als Arbeitspakete identifiziert wurden: zu denen Herr Heuvelmann in Vertretung der Gruppe in die MARC/RDA Working Group Sitzungen gehen möchte:

- Content type, media type, carrier type
- URIs
- neue Elemente im Authority Format

Herr Heuvelmann nimmt die Überlegungen der EG Datenformate in die RDA/MARC Working Group mit. Es wird darum gebeten, der EG Datenformate regelmäßig zu berichten, wenn Überlegungen einer Weiterentwicklung oder Abweichungen von MARC in der Regelwerk/Datenformate Community im Gespräch sind.

Er bittet seinerseits die Mitglieder darum, in den Institutionen an den Stellungnahmeverfahren zu den RDA mitzuwirken, und zwar besonders aus der MARC-Perspektive.

c. Mehrbändig begrenzte Werke

Ein Treffen der ALEPH-Verbände mit der Firma ExLibris ist für Januar geplant. Es sollen dabei die Rahmenbedingungen (zeitliche und ökonomische Faktoren) des Systemangebots bezüglich des Intern-Formats besprochen werden. Es wird klar gestellt, dass das im September vorgelegte Diskussionspapier für die Konzipierung des Internformats und nicht für das vom Austausch betroffene Format gilt. In den Verhandlungen mit ExLibris wird angestrebt die Festlegungen der EG Datenformate zu mehrbändigen Werken beizubehalten.

Die ALEPH-Verbände werden darum gebeten, zu prüfen, wo der Datenaustausch von der Vereinfachung des Modells betroffen ist, und entsprechend zu berichten. Die Vorteile der Austauschbarkeit vs. der Katalogisierungsergonomie sollten in Betracht gezogen werden. Es wird daran erinnert, dass die DFG eine einheitliche Anwendung des Formats anstrebt – eine finanzielle Hilfe dazu liegt im Bereich des Möglichen.

d. Abbildung von RSWK-Daten in MAB2 und MARC 21

Zum Tagesordnungspunkt 5.d liegt den Teilnehmern das Papier V-080 "Abbildung von Schlagwörtern nach den RSWK in MARC 21 in Datenlieferungen der Deutschen Nationalbibliothek" vor.

Herr Heuvelmann berichtet, dass das Papier V-080 "Abbildung von Schlagwörtern nach den RSWK in MARC 21 in Datenlieferungen der Deutschen Nationalbibliothek" akzeptiert wurde. Ein Papier zur Konversion aus MAB nach MARC soll noch veröffentlicht werden.

Es wird in der Gruppe hinterfragt, warum man beabsichtigt, zusätzlich SWD-Nummern mit ISIL "0?599c" zu liefern, die bisher in MAB nicht geliefert werden konnte und wurde. Es wird angemerkt, dass auch in diesem Punkt, die nach MARC vorgesehenen Möglichkeiten genutzt werden sollen. Die Aleph-Verbände liefern nur eine Nummer.

e. Abbildung von DDC-Daten in MAB2 und MARC 21

Zu diesem Tagesordnungspunkt wird berichtet sobald es Neuigkeiten gibt.

f. MARC-Antrag 2008-02

Zum Tagesordnungspunkt 5.f liegen den Teilnehmern die Papiere V-079 "Abbildung der ISSN-L in MARC 21 und MAB2" und MARC_2008-02 "MARC-Änderungsantrag: Festlegung für die MARC-21-Anwenderebene" vor.

Die Einführung eines anwenderspezifischen Feldes zum Transport eines Produktsigels wurde von der ZDB für "virtuelle Sammlungen" in MARC-Feld 950 festgelegt. Zwischenzeitlich wurden die Festlegungen zur Verteilung der nationalen, regionalen und lokalen MARC-Felder getroffen; dabei wurde das anwenderspezifische MARC-Feld 950 dem Bereich des GBV zugeteilt. Es wird vorgeschlagen, das erste Feld aus dem national zu definierenden Bereich, nämlich Feld 912, für den Transport der Produktsigel zu verwenden. Es wird diskutiert, ob das hier beantragte Feld 912 stärker entsprechend der benötigten Angaben mit wiederholbaren Unterfeldern strukturiert werden soll. Es wird beschlossen, das Feld 912 für die ZDB wiederholbar zu machen, die Unterfelder jedoch nicht wiederholbar zu machen.

g. ISSN-L in MAB2 und MARC 21

Zum Tagesordnungspunkt 5.g. liegt den Teilnehmern das Papier V-075a "Entwurf für ein MAB-Feld im Segment "Anwenderspezifische Angaben" zur vollständigen Abbildung von MARC-21-Feldern" vor.

Zum konversions- und verlustfreien "Parken" von MARC-21-Feldern in MAB2 ist ein Feld anzulegen. Es stellt lediglich eine "Hülle" dar, um die Angaben zum MARC-21-Teilformat, zu der Feldnummer, den Indikatoren und den Unterfeldern in MAB2 abzubilden. In der 13. Sitzung der Expertengruppe Datenformate am 16. September 2008 wurde ein erweiterter anwenderspezifischer Bereich mit den Feldern 092 bis 099 vereinbart und für das Feld die Feldnummer 099 gewählt. Das Feld kann sowohl für einzelne Datenelemente als auch für ganze Datensätze verwendet werden, wie es gebraucht wird ist Entscheidung der Anwender. Das Feld wird demnächst in ZDB-Daten mit der ISSN-L auftauchen. Dazu wird es eine Ankündigung geben.

Es wird vorgeschlagen, dass das Beispiel im Papier geändert und die ISSN-L im Parkfeld gezeigt werden sollen.

h. IDN-Transport MARC 21

Für die neu aufgebauten MARC-21-Datenlieferungen der Deutschen Nationalbibliothek ist eine Änderung gegenüber den MAB2-Lieferungen angekündigt und diskutiert worden: Die Belegung des Feldes 001 "Control number" erfolgt mit der DNB-internen Identifikationsnummer des Datensatzes. In MAB2 sind teilweise andere Nummern (GKD-Normnummer; SWD-Normnummer; ZDB-Nummer) anstatt der Systemnummer in 001 geliefert worden. Nach MARC 21 muss die interne Nummer in 001 stehen, in 003 der MARC Organization Code der liefernden Institution. Weitere Nummern, wie Norm- und ZDB-Nummern, sowie regionale IDNs, werden in anderen MARC-21-Feldern (016 bzw. 035) geliefert.

Die Verknüpfung mit Hilfe von Identifikationsnummern ist in MARC 21 breiter angelegt als in MAB2: Die entsprechenden Unterfelder \$0 (für eine Verknüpfung hin zu einem Normdatensatz) und \$w (für eine Verknüpfung hin zu einem Titeldatensatz) sind wiederholbar, es können also parallel Verknüpfungen zu demselben Zieldatensatz mit Hilfe von zwei verschiedenen IDNs des Titeldatensatzes angegeben werden. Die IDNs sind jeweils mit einem vorgestellten MARC Organization Code gekennzeichnet. Inkonsistenzen, wie teilweise in MAB2 notwendig, können so vermieden werden.

Auch die Verknüpfung von einem Holdings-Datensatz hin zum dazugehörigen Titel-Datensatz erfolgt in doppelter Form: Einerseits wird die DNB-interne Systemnummer des Titeldatensatzes im Feld 004 geliefert, und parallel die ZDB-Nummer des Titeldatensatzes im Feld 016.

Diese neue Belegungspraxis wird diskutiert. Ein gewisser Mehraufwand beim Aufbau der MARC-21-Importschnittstellen ist damit verbunden; Feld 001 dient unabhängig von der Art des Formats als Primärschlüssel. Es wird zugesagt, die Auswirkungen auf die Systeme zu prüfen.

i. Dokumentation zur Anwenderenebene

Dieser zusätzliche Tagesordnungspunkt klärt die Frage der Beschreibung von anwenderspezifischen MARC-Feldern. Diese Felder werden in einem Dokument auf Deutsch "Beschreibung der MARC-21-Formatelemente in den Titeldaten" definiert und dokumentiert und wurden an Frau Behrens-Neumann für die Veröffentlichung durch die AG Verbund abgegeben. Weitere Details werden von Frau Block und Frau Oehlschläger nach der EG Datenformate Sitzung besprochen.

Das Editionssystem Polly wird über die Felddescription hinaus eine vollständige Übersetzung von MARC 21 Concise bieten. Es gibt keine spezifischen Termine, es handelt sich um ein laufendes Vorhaben.

j. Weitere Planungen

Die nächste Sitzung ist für den 1. April 2009 geplant. Es besteht die Option die Sitzung abzusagen wenn nicht genug Stoff vorhanden ist.

6. Verschiedenes

Zu diesem Tagesordnungspunkt gibt es nichts zu berichten.