

**Kommission für Erschließung und Metadaten
des Bibliotheksverbundes Bayern (KEM)
Arbeitsgruppe Sacherschließung (AGSE)**

**Protokoll der 10. Sitzung der Amtsperiode 2014 – 2018 am 05.03.2018,
Bayerische Staatsbibliothek München, 10.30 - 15.00 Uhr**

Teilnehmer:

- Dr. Gerhard Stumpf, UB Augsburg (Vorsitz, Moderation)
- Edith Dilber, BVB-Verbundzentrale (Gast zu TOP 10)
- Eva-Maria Gulder, BSB München
- Werner Holbach, BSB München
- Dr. Sonja Kobold, Bibliotheca Hertziana Rom
- Tatjana Merda, Bibliothek der TH Nürnberg
- Dr. Gabriele Schweikl, UB Regensburg (Protokoll)
- Peter Thiessen, BVB-Verbundzentrale (Gast)
- Monika Vogler, UB Augsburg
- Barbara Wolf-Dahm, UB Augsburg (Gast)

Entschuldigt:

- Michael Franke-Maier, UB der FU Berlin
- Christian Rüter, UB der HU Berlin

TOP 1: Begrüßung, Formalia, Tagesordnung

Zur vorab verschickten Tagesordnung werden folgende Punkte ergänzt:

- TOP 15: Umsetzung der Gattungsbegriffe „gatbeg“
- TOP 16: Formatänderungen der DNB

Wegen Anwesenheit von Frau Dilber wird TOP 9 vorgezogen.

TOP 2: Berichte aus den Gremien

a. Kommission für Erschließung und Metadaten:

Nacharbeiten aus der Prä-RDA-Umsetzung können noch bis zum 15.3. erfolgen, ab dem 16.3. werden dann Gesamtabzüge an die Bibliotheken ausgeliefert bzw. eingespielt. Der Gesamtabzug der ZDB wird ausgewertet und geprüft, ob wesentliche Umsetzungen fehlen. Insgesamt wurden 93% der B3Kat-Sätze im Prä-RDA-Projekt angefasst.

Der unerwünschte Import der maschinell ermittelten Schlagwörter der DNB über HeBIS (Z39.50) wurde mittlerweile gestoppt.

Lokalsysteme: Ein Abschlussmeeting zum Thema „Test Alma und Open Source Systeme“ findet Mitte März statt. Auch wenn es positive Ergebnisse beim Alma Test gibt (Einbindung der e-Medien, Erfassung von GND-Sätzen), so sind doch auch noch deutliche Defizite feststellbar.

Die nächste KEM-Sitzung findet am 17.7. statt, die Sitzung mit den Lokalredaktionen am 16.10.

b. Fachgruppe Erschließung

Schulungen zum 3R-DACH-Projekt können frühestens ab Herbst 2019 stattfinden. Nachdem die Struktur vom RSC aufgebaut wurde, könnten jetzt Texte einfließen und von Seiten des RSC mit der Revision begonnen werden. Erste Texte in einer Beta-Version sollen zum 13.6.2018 zur Verfügung stehen, mit anschließenden Updates. Diskutiert wurde die Zusammenarbeit mit und Abgrenzung zum GND-Ausschuss – alles, was regelwerksrelevant ist, sollte nach Meinung der Fachgruppe auch dort beschlossen werden. Speziellere Themen der Besprechungen waren u.a.: Reproduktionen (hierzu wird ein Expertenteam eingesetzt) und die Normierung bei Klöstern Stiften etc. (das Thema wird bis zum Erscheinen des neuen toolkits vertagt).

Mit der Einführung eines JIRA-Ticketsystems für Fragen, wie es bereits für die GND verwendet wird, ist die Fachgruppe Erschließung einverstanden.

c. GND-Ausschuss:

Der GND-Ausschuss hat sich mit zwei großen Themen befasst, für die eigene Expertengruppen bestimmt wurden: Namensdatensätze (Tn) und Werke. Das Expertenteam Tn empfiehlt dem GND-Ausschuss, auf Tn-Datensätze als Teil der GND zu verzichten. Zur Fragestellung Werke und deren Expressionen sollen Experten aus den Arbeitsgruppen Alte Drucke, Bild, Musik, Kunst und Handschriften hinzugezogen werden. Die Ansetzung der Expressionen biblischer Werke wird vertagt, bis die neue RDA-Version vorliegt.

Das Entwicklungsprogramm der GND wird u.a. in der Konferenz GNDCon am 3.-4-12.2018 in Frankfurt vorgestellt (GND4C für Kulturdaten, ORCID, Weiterentwicklung der GND etc). Für die Verwendung von Codes für familiäre Beziehungen wird ein Formatvorschlag erstellt (Bedarf hierzu besteht bei den Anwendern des GND-Webformulars), was aber nicht als dringlich eingestuft wird.

Zu geografischen Datensätzen hat ein Workshop stattgefunden, zu dem viele Regionalbibliografien eingeladen wurden. Es wurden konkrete Wünsche geäußert, ebenso wurde vorgeschlagen, auf außenstehende Systeme zu verlinken - ähnlich wie bei Bildern, wo die Verlinkung zu IconClass erfolgt.

d. Expertenteam „RDA-Anwendungsprofil für die inhaltliche Erschließung:

Das Expertenteam beschäftigt sich mit der Frage, ob mit den bisherigen RSWK die Grundanforderungen an ein Regelwerk erfüllt werden hinsichtlich Qualität, Vollständigkeit und Themenbildung. Das betrifft auch die Problematik der Verwendung von engen oder weiten Schlagworten und damit die Genauigkeit der Erschließung. Anhand einer vorliegenden Themen- und Prioritätenliste werden Prioritäten festgelegt, Schwerpunkte werden sein: Werkdatensätze, Zeitangaben, Aspekt Rollen-Operatoren (zb. Motiv-Schlagworte).

Zur Frage nach dem Erschließungskonzept der DNB wird der Workshop für die Mitglieder des Standardisierungsausschusses am 9.3. abgewartet.

Bericht vom Workshop Maschinelle Inhaltserschließung am 15.12.2017 (Hr. Stumpf, Hr. Holbach)

- Maschinelle Schlagwortvergabe der DNB:

Die DNB verwendet als Wörterbuchbasis alle GND-Sätze mit Teilbestandskennzeichen s (Schlagwörter). Die Reihe O wird seit 2017 komplett bearbeitet (ca. 350 Dokumente pro Tag); alle Jahrgänge 2010-13 sollen noch nachgearbeitet werden. Die zu scannenden Inhaltsverzeichnisse der Publikationen aus Reihe B und H werden vom Scanpersonal intellektuell ausgewählt. Jeder Bibliografiejahrgang wird komplett wiederholt bearbeitet.

Rahmenbedingungen und Prozesse wurden an Hand umfangreicher Präsentationen beschrieben. Es handelt sich nicht um ein „lernendes“ System, da die Ergebnisse früherer Bearbeitungsfälle nicht implizit für die Verbesserung herangezogen werden. Verbesserung lässt sich nur durch die (intellektuelle) Ergänzung der Wörterbücher erreichen. Der Konfidenzwert ist ein rein maschinelles Kriterium (interner Schwellenwert), der u.a. von der Textgröße abhängt und als absoluter Qualitätswert nicht taugt.

Ein Terminologiemanagementsystem (TeMa) soll künftig Vorschläge für die Ergänzung der GND bereitstellen. Das Projekt MAEN zielt auf die ergänzende Auswertung englischsprachiger LCSH-Termini ab.

Geplant ist der Einsatz einer Textanalysesoftware, die bestimmte Textteile erkennt und weniger brauchbare Passagen ggf. ausscheidet.

RDA-Formbegriffe werden generell nicht vergeben.

Zwischen Parallelausgaben der Reihe A (Print – E-Book) wird zu internen Untersuchungszwecken die intellektuelle bzw. maschinelle Inhaltserschließung ausgetauscht bzw. ergänzt.

Über den Einsatz eines Vorschlagssystems wird an der DNB nachgedacht.

- Maschinelle Vergabe von DDC-Kurznotationen (bisher Sachgruppen):

Hier handelt es sich um ein „lernendes“ Verfahren. Bisher wurden Kurznotationen für Medizin (140 Klassen) und Informatik (72 Klassen) vergeben. Das wird auf alles ausgeweitet. Die Benennungen der Kurznotationen sollen ebenfalls ausgegeben bzw. daraus GND-Termini erzeugt werden.

In der Diskussion wurde angeregt, die Inhaltserschließung mit einem Werkclustering zu verbinden. Die DNB nahm Anregungen hinsichtlich Kooperation, besserem Austausch und konkreterer Information auf.

TOP 4: PRDA-Projekt: Formschlagwörter / Formangaben

a. Nacharbeiten

Für die Nacharbeiten der PRDA-Umsetzung im Bereich der Formschlagwörter und Formangaben werden folgende Vorgaben beschlossen:

- i. Löschen von 064a Biografie, wenn 064a Autobiografie bereits vorhanden ist,
- ii. Löschen folgender Formschlagwörter in 9xx: Videokassette, Mikroform, Medienkombination, Elektronische Publikation. Das Schlagwort f Geschichte in 9xx soll in z umgesetzt werden. [*Beide Aufträge wurden am 13.04.2018 von der Verbundzentrale erledigt.*]

Die Permutationsmuster sollen bei Löschungen komplett gelöscht werden..

Herr Thiessen prüft außerdem, ob eine Belegung der IMD-Felder in Abhängigkeit von ASEQ 433 möglich ist.

Der Rest der Liste „Restliche f in Schlagwortfolgen“ soll kooperativ intellektuell korrigiert werden.

- b.** Einhaltung der Erfassungsregeln im B3Kat / Fremddatenimporte (Headings-Index SWW nicht neu aufgebaut)

Bei der chronologischen Sortierung zeigt sich, dass hauptsächlich über Fremddateneinspielungen noch Form-Schlagwörter in Aleph aufgenommen werden. Die mit ZDB-Titeln verknüpften Form-Schlagwörter wie Verzeichnis, Adressbuch, Zeitschrift u.ä. müssten in der ZDB gelöscht werden.

TOP 5: Bereinigung des Feldes ASEQ 720

Auf der Basis einer überarbeiteten Negativliste von Feldinhalten oder Feldanfängen, die Herr Stumpf erstellt, soll eine Testbereinigung von ASEQ 720 vorgenommen werden.

TOP 6. Dublette Schlagwortfolgen (902 / 907 ...): Bereinigungsauftrag

Der Auftrag SE15 (Bereinigung dubletter Schlagwortfolgen in 9xx) wird den Vorgaben gemäß (siehe Protokoll der letzten AGSE-Sitzung) noch vor dem 15.3. durchgeführt werden, die Versorgung der Lokalsysteme dann erst nach diesem Termin. Betroffen sind ca. 175.000 Datensätze. [*wurde erledigt*]

TOP 7. Slash in DDC-Notationen

Die uneinheitliche Verwendung des Slashes in DDC-Notationen macht in den angeschlossenen Systemen (wie Primo oder WebdeweySearch) Probleme. Dabei zeigt der Slash an, wo die Kurznotationen der abridged edition enden, eventuell auch die Version.

Herr Thiessen wird als Beispielgrundlage für eine Entscheidung über eine mögliche Datenbereinigung mit Anpassen des Importvorgangs eine Analyse der Fälle vornehmen, in denen der Slash in DDC-Notationen vorkommt.

TOP 8. Erweiterte Schlagwort-Nachführung

Herr Thiessen berichtet, dass die Programmiererweiterung zur Anreicherung des ASEQ-feldes 064 (sofern es nicht bereits mit gleichem Inhalt belegt ist) in der Entwicklung durch Herrn Lauer ist. Details siehe Protokoll der letzten AGSE-Sitzung.

TOP 9: Einrichtung einer Thesaurus-Library in Aleph

Herr Thiessen berichtet vom Stand des GNOMON-Projektes zur Verwaltung einer Thesaurus-Library in Aleph:

- Im Interformat ist im Titelbereich das ASEQ-Feld 711 mit Unterfeld \$9 im Schulungssystem eingerichtet.

- Für Endnutzer wird es eine Touchpoint-View geben, bei der in den Titeln die ID-Nummer für die Verknüpfung zum Normdatensatz mitgeliefert wird.
- Die ID-Nummer sieht wie die BV-Nummer aus, beginnend mit TH (wie Thesaurus), der ISIL für den Gnomon Thesaurus und einer 9stelligen Nummer.
- Die Datensätze werden löschgeschützt sein.
- Ab 9. April werden neue Titel im B3Kat erfasst.

TOP 10: RVK

Frau Dilber stellt kurz den von ihr erstellten Formatvorschlag für das neue Feld ASEQ 701 für die Erfassung von Normdatei-kontrollierten RVK-Notationen vor.

- Unklar ist die Frage, ob \$2 neben \$9 benötigt wird. In \$9 wird angezeigt, aus welcher Quelle die Notation kommt. Mit der entsprechenden Belegung von \$2 kann das Feld 701 auch für andere Systematiken verwendet werden, aber auch bei Bedarf erst später aktiviert werden. Frau Dilber wird klären, ob \$2 notwendig ist.
- Die derzeit in 700g abgelegten Notationen sollen auf Vorhandensein in der Normdatei BVB13 überprüft werden. Normdateikonforme Einträge werden in das Feld 701 überführt und dann in 700g gelöscht. Als Resultat bleiben nur noch nicht normierte Notationen in 700g („lokale Anwendungen“ oder Zeitschriften-„Notationen“).
- Eine grüne Fehlermeldung beim Vorhandensein einer Notation in 700g soll auf Feld 701 hinweisen.
- Der überregionale Austausch normierter Daten wird über MARC 084 vorgenommen.

Beschluss: Das neue ASEQ Feld 701 soll mit der vorgeschlagenen Feldstruktur eingeführt werden.

Top 11: Gateway Bayern: Darstellung der GND-Schlagwörter

Eine durch Herrn Stumpf erstellte Liste der MARC-Felder, die für die Darstellung der GND-Schlagwörter relevant sind, liegt mittlerweile vor. Frau Wolf-Dahm wendet ein, dass eine einheitliche Übernahme der Daten für alle unterschiedlichen OPACs möglich gemacht werden sollte. Zu beachten seien deshalb die Vorgaben aus RDA (Anhang E1.1 Darstellung von Sucheinstiegen) mit der Zuordnung von RDA-Elementen zu Ansetzungsformen und –verweisungen für Personen, Körperschaften, Einheitssachtitel und Titelverweisungen. Herr Stumpf will sich seine Zusammenstellung vor dem Hintergrund dieser RDA-Vorgaben noch einmal ansehen.

TOP 12: GND: Sachstandsbericht GND-Kooperation und Datenpflege

- Im Hintergrund läuft viel Datenpflege samt Dublettenbereinigung. Dringend nötige Korrekturen werden auf Zuruf erledigt
- Punktuelle Umarbeitungen: „Diözese“, „Katholische Kirche“ u.a. - „Universitäten“ sollte von der DNB aufgegriffen werden
- Für juristische Werke wurde eine Erfassungshilfe erstellt und mit Aleph-Beispielen ergänzt.

- Im Zuge des Projektes der Bayerischen Staatsbibliothek "Normdaten im Projekt Erschließung und Digitalisierung von Einbänden als selbständige Kunstobjekte" werden etwa 200 neue Sachbegriffe zur Einbandkunde vom BVB (UB Augsburg) in die GND eingebracht. Sie werden im o.g. Projekt zur Verknüpfung mit der GND benötigt.

TOP 13: Umsetzung der Zeitschlagwörter in Zeitcodes im Kontext der Diskussion über die Erschließung des Zeitaspekts

Eine durch Herrn Stumpf erstellte „Vorgabe für die Vergabe eines Zeitcodes für den thematisierten Zeitaspekt in B3Kat“ liegt vor, daneben wird als Arbeitsgrundlage des Expertenteams Sacherschließung auch die Masterarbeit von Michael Wening („Die Zeitfacette in bibliothekarischen Erschließungen“) verwendet.

Da das entsprechende Feld z. B. im Gateway Bayern auch ausgewertet werden muss, bräuchte man auf der Nutzeroberfläche eine Funktionalität, mit der Zeitangaben gefiltert werden können. Nur wenn diese definierbar ist, macht die Einführung des Zeitcodes Sinn. Herr Thiessen wird der Frage nachgehen, der TOP wird damit auf die nächste Sitzung vertagt.

TOP 14: Einspeicherung von Sacherschließungselementen aus FID-Bibliographieprojekten der BSB: Stand der Projekte

DHB Deutsche Historische Bibliografie:

- Bei der Beantragung der MARC-Source-Codes ist es wegen Rückfragen seitens der DNB zur langfristigen URL der Klassifikationssysteme zu Verzögerungen gekommen. Mit der Einspeicherung von Sacherschließungsdaten wird noch gewartet, bis der MARC-Code endgültig genehmigt ist. Dann werden im Rahmen der DHB im Fall der JDG-Altdaten prototypisch folgende Sacherschließungselemente in den B3kat eingespeichert:

711_ = Alte JDG-Schlagwörter (als Schlagwortfolge: SW / SW / SW); 711_ \$x dhb-jdg (als MARC Source Code gemeldet)

711_ = Verbalisierung der JDG-Klassifikation (Syntax: Oberklassenbezeichnung/_/...
 _/_Klassenbezeichnung); 711_ \$x dhb-jdgklass/1, dhb-jdgklass/2 (als MARC Source Code gemeldet)

700_ = JDG-Klassifikation; \$2 dhb-jdgklass/1, dhb-jdgklass/2

- Auch für weitere Datenbestände mit eigenen Klassifikationssystemen, z.B. IfZ und Bayerische Bibliographie, ist im Rahmen der DHB ein analoges Vorgehen geplant.

Aufsatzerschließung / Projekt mit AvantgardeLabs:

- Die Ergebnisse der Tests zur Anreicherung mit Wikipedia-Konzepten für Geschichte/Osteuropa über AvantgardeLabs waren bisher nicht zufriedenstellend, sodass Nachbesserungen notwendig sein werden.

TOP 8: Formatänderungen DNB

Aus den mit Rundschreiben vom 12.02.2018

(http://www.dnb.de/SharedDocs/Downloads/DE/DNB/service/rundschreiben20180212AenderungenTiteldatenMarc21.pdf?__blob=publicationFile) von der DNB angekündigten Formatänderungen, die zum 15.05.2018 in Kraft treten sollen, ergeben sich für die Sacherschließung im B3Kat verschiedene Fragen:

- MARC-Feld 650 (Nebeneintragung unter einem Schlagwort, Sachschlagwort):

„Mehrere Schlagwörter aus Fremddaten ohne Verknüpfung zur GND, die in einem Unterfeld \$a ausgeliefert werden, werden nun in der Regel durch „;_“ voneinander getrennt, wenn die Struktur der Quelldaten dies ermöglicht“. Hierzu stellt sich die Frage, wie die Daten übernommen werden sollen, bisher üblich sind „_/_“ (Slash) und „_–“ (Bindestrich).

Herr Thiessen wird zunächst zusammenstellen, was über Feld 650 in Aleph kommt.

- MARC-Felder 082 (Notation nach der Dewey Decimal Classification) (NR) und 083 (Zusätzliche DDC-Notation) (R): Auslieferung maschinell erstellter DDC-Kurznotationen und entsprechende Ergänzung zu den Provenienzinformatoren im verknüpften MARC-Feld 8832)

Die DDC-Kurznotationen werden in \$2 als kDNB gekennzeichnet und eingespielt, eine Unterscheidung von maschinell und intellektuell erstellten Kurznotationen erfolgt nicht mehr. Eine AG der KEM befasst sich mit dem Arbeitsauftrag „Übernahme maschinell erstellter SW“ und wird in diesem Kontext auch das Thema maschinelle Notationen bzw. Kurznotationen klären. Aus Sicht der AGSE spricht nichts gegen eine Übernahme der maschinell erzeugten Kurznotationen, wenn sie als solche gekennzeichnet sind.

- Synthetische Notationen: Zusammensetzung der DDC-Notationen und Einspielung in MARC 085

Aufgrund der gestiegenen Verwendung der \$8-Verknüpfungsfelder in den DNB-Daten in MARC 085 (mit Informationen zur Sequenzierung und Zuordnung) gestaltet sich die Verarbeitung u.a. der gelieferten DDC-Notationen im Konverter sehr aufwändig und schwierig, was die Zusammensetzung der lieferten Notationen aus den gelieferten Bestandteilen anbelangt. Die Daten werden im B3Kat in ASEQ 705 \$a \$e und \$f abgelegt. Da zusätzlich noch die vollständige Notation geliefert wird, stellt sich die Frage, ob nicht auf die Zusammensetzung der synthetischen DDC-Notationen verzichtet werden kann und stattdessen die in MARC 082 und MARC 083 gelieferten Notationen verwendet werden können.

Herr Thiessen wird die Thematik noch einmal aufbereiten.

TOP 16: Gattungsbegriffe „gat-beg“

Inhalte in ASEQ 710f, die mit „gat-beg“ gekennzeichnet sind, sollen in ASEQ 711f transferiert werden. Derzeit werden die Daten unstrukturiert abgelegt. Herr Thiessen wird eine Analyse machen. Der Konverter muss entsprechend angepasst werden.

TOP 17: Planung für die nächste Sitzung am 12.06.2018

Die nächste Sitzung am 12.06.2018 wird in Berlin etwas verkürzt stattfinden, da der Bibliothekartag mit einer Session zum Thema Erschließung bereits um 13 Uhr beginnt. Nächste KEM-Sitzung: 17.07.2018