

Expertengruppe Sacherschließung

## Protokoll der 10. Sitzung - Status: Verabschiedet

14. März 2011, 10.30 Uhr bis 17.00 Uhr

Deutsche Nationalbibliothek Frankfurt am Main

### Teilnehmerinnen und Teilnehmer

Susanne Bohner	SWB (BSZ Konstanz)
Dorothea Büning	HBZ Köln
Dörte Braune-Egloff	KOBV (UB FU Berlin)
Urs Frei	Schweizerische NB Bern
Werner Holbach	BSB München
Dr. Volker Schümmer	Kunstabibliotheken-Fachverbund Florenz-München-Rom (ZIKG München)
Armin Kühn	SWB (BSZ Konstanz)
Bettina Kunz	GBV (SUB Göttingen)
Esther Scheven (Vorsitz)	DNB Frankfurt am Main
Dr. Gerhard Stumpf	BVB (UB Augsburg)
Dr. Joachim Eberhardt	HBZ (Lib Detmold)
Prof. Heidrun Wiesenmüller	SWB (HdM Stuttgart)
Christa Wittrock	HeBIS (UB Kassel)

### Entschuldigt

Petra Friedmann	ekz
Almut Ilsen	SB zu Berlin
Heike Krems	SB zu Berlin
Dr. Ulrike Reiner	GBV (VZG Göttingen)
Dr. Margit Sandner	OBV (UB Wien)
Vera Uhlmann	Schweizerische NB Bern

### Gäste

Dr. Christian Baumann	DNB Frankfurt am Main
Brigitte Bernhard	DNB Frankfurt am Main
Reinhold Heuvelmann (ab TOP 3.2.2)	DNB Frankfurt am Main
Ulrike Junger	DNB Frankfurt am Main
Dr. Julijana Nadj-Guttandin	DNB Frankfurt am Main
Barbara Pfeifer	DNB Frankfurt am Main

### Protokoll

Karen Köhn	DNB Frankfurt am Main
------------	-----------------------

# Tagesordnung

1. Begrüßung, Genehmigung der Tagesordnung
2. Verabschiedung des Protokolls der Sitzung am 2.12.2010
3. GND
  - 3.1. Stand und Zeitplan
  - 3.2. Spezielle Themen der Sacherschließung
    - 3.2.1 Übergangsregeln  
(Mehrgliedrigkeit, Ländercode, Werke der Musik)
    - 3.2.2. Splits bei Gebietskörperschaften
    - 3.2.3. Mehrgliedrige Oberbegriffe
    - 3.2.4. Umsetzung von Mundarten
    - 3.2.5. Liste der \$4-Codierungen
    - 3.2.6. Umsetzung von Hinweissätzen
    - 3.2.7. Match-and-Merge
    - 3.2.8. Geografikum nach Formschlagwörtern (z. B. Kongress)
4. Automatische Erschließung für Netzpublikationen (DNB-Projekt)
5. Verschiedenes
  - 5.1 Rangfolge der Nachschlagewerke
  - 5.2 RSWK-Änderungen
  - 5.3 Umfragebogen
  - 5.4. Retrieval-Möglichkeiten
  - 5.5 Sachgruppen in MAB2-Titel

## Tischvorlagen / Unterlagen

befinden sich im Wiki unter <https://wiki.d-nb.de/x/KIGRAg>

- |           |  |
|-----------|--|
|           | Tagesordnung   |
| TOP 3.1   | Meilensteine-Zeitplan März 2011  |
| TOP 3.2.1 | GND-Übergangsregeln allgemein ( <a href="https://wiki.d-nb.de/x/7oSRAg">https://wiki.d-nb.de/x/7oSRAg</a> )  |
| TOP 3.2.1 | Übergangsregeln RSWK-Tabelle   |
| TOP 3.2.1 | Übergangsregeln gio, giw   |
| TOP 3.2.1 | Übergangsregeln Ländercode   |
| TOP 3.2.1 | Übergangsregel2 Abkürzungen Instrumente  |
| TOP 3.2.2 | Splits bei Gebietskörperschaften (Wiesenmueller)   |
| TOP 3.2.2 | Splits bei Gebietskörperschaften Stellungnahmen  |
| TOP 3.2.2 | Stellungnahme DNB zum Papier von Frau Wiesenmueller  |
| TOP 3.2.3 | MO Kommentar, Umfrage  |
| TOP 3.2.4 | Mundartenumsetzung GND   |
| TOP 3.2.4 | Auflösung 845 Mundarten Sprachen   |
| TOP 3.2.5 | Änderungsvorschläge zur Codeliste Relationstypen<br>(vgl. <a href="https://wiki.d-nb.de/x/Xx2YAg">https://wiki.d-nb.de/x/Xx2YAg</a> , 9. Codelisten) |
| TOP 3.2.6 | Hinweissätze   |
| TOP 3.2.7 | Match-and-Merge  |
| TOP 3.2.8 | Geografikum nach Formschlagwörtern   |
| TOP 4     | Petrus   |
| TOP 5     | Vorschlag zur Änderung RSWK (Sprachen, Ethnographika)  |

## 1. Begrüßung, Genehmigung der Tagesordnung

Zur neuen Periode der Expertengruppe begrüßt Frau Scheven die Anwesenden. Nach einer Vorstellungsrunde wird Frau Prof. Heidrun Wiesenmüller als stellvertretende Vorsitzende einstimmig durch die Anwesenden bestätigt.

Herr Dr. Stumpf bittet, das Informationsverfahren bei Ansetzungsänderungen als Tagesordnungspunkt unter Verschiedenes in die Tagesordnung aufzunehmen. Da sich die EG Normdaten und die AG Datenformate am 11. April 2011 in der DNB zu einer gemeinsamen Sitzung zum Thema GND-Redaktionsverfahren treffen werden, wird der Punkt nicht in die Tagesordnung aufgenommen. Die vorliegende Tagesordnung wird genehmigt.

## 2. Verabschiedung des Protokolls der Sitzung am 2.12.2010

Die vorliegenden Änderungswünsche von Frau Braune-Egloff und von Frau Dr. Sandner werden in das Protokoll der letzten Sitzung übernommen. Das Protokoll wird verabschiedet.

## 3. GND

### 3.1 Stand und Zeitplan

Der am 3.12.2010 in der SWD-Redaktionsgruppe vorgestellte Zeitplan wurde leicht verändert. Frau Scheven erläutert die Tischvorlage „Meilensteine-Zeitplan März 2011“.

Ende Januar fand in einem ersten Schritt die Migration aller Normdaten im GND-Testsystem statt, in einem zweiten Schritt erfolgt zurzeit die Relationierung, die zum 15. März 2011 fertiggestellt wird. Zum 15. März wird auch eine erste Test-Gesamtlieferung zur Abholung bereitgestellt.

Viele inhaltliche Vorgaben konnten in dieser ersten Stufe noch nicht programmiert und umgesetzt werden. Die Programmierung der zweiten Stufe hat begonnen; ihre Umsetzung, Migration und Relationierung soll im Juni erfolgen.

Da zum GND-Testsystem noch nicht alle Anwesenden Zugriff haben, sollen entsprechende Zugangsinformationen seitens der DNB zur Einrichtung der Zugänge in den Verbänden bereitgestellt werden. *[zwischenzeitlich per E-Mail erfolgt]*

Die MARC 21-Onlineschnittstelle für die GND-Onlinekommunikation wurde Anfang Januar einem ersten Funktionstest mit Testdaten unterzogen, zurzeit finden Partnertests der Stufe 1 statt. Partnertests der Stufe 2 sind im Juli geplant.

DNB beschäftigt sich jetzt schon mit den Vorarbeiten für das Redaktionsverfahren; Festlegungen sollen auf der gemeinsamen Sitzung der EG Normdaten und der AG Datenformate am 11. April 2011 in der DNB getroffen werden. Die Aleph-Verbände müssen ihre Befugnis-Verwaltung in Abstimmung mit den Pica-Verbänden entwickeln, da sie nur mit jeweils einer Kennung über die ONS auf die Master-GND zugreifen. Ende September soll die Redaktionsrichtlinie fertiggestellt sein.

Das Match-and-Merge-Verfahren für doppelte Entitäten in GKD und SWD soll gemäß mehrheitlichen Beschluss erst nach Produktivschaltung der GND in 2012 erfolgen. Allerdings müssen technische Abläufe DNB-intern getestet werden. Das ist für Ende Mai 2011 geplant. DNB sagt die Ergänzung des fehlenden Zeitpunktes zur endgültigen Zusammenführung im vorgelegten Zeitplan zu.

Der GND-MARC-Formatentwurf 0.7 wurde am Sitzungstag versendet. Der GND-Formatentwurf 1.0 einschließlich Originalschriftlichkeit ist für Ende März geplant, wird voraussichtlich aber erst nach der gemeinsamen Sitzung der EG Normdaten und der AG Datenformate am 11. April 2011 freigegeben werden. *[zwischenzeitlich erfolgt]*

Auf der Sitzung der AG Datenformate am 17.2.2011 wurden die beantragten Datenfelder für die GND mit Ausnahme von Feld 669 für Benutzungshinweise genehmigt. Für Benutzungshinweise wurde das offizielle MARC-Feld 680 vorgeschlagen.

Am 10.3.2011 fand in der DNB ein Treffen der Pica-Anwender statt; ein Treffen der Aleph-Anwender wird am 23.3. in Köln stattfinden.

Das Schulungskonzept soll in der Sitzung im April vorgestellt werden. Es wird vorgeschlagen, Schulungsunterlagen arbeitsteilig, evtl. in einer dafür zu bildenden Arbeitsgruppe, zu erstellen. Insbesondere Schulungsunterlagen zu den Übergangsregeln können einheitlich für alle Anwender erstellt werden.

Die Schulungen sollen in zeitlicher Nähe zur Produktivschaltung der GND erfolgen. Der Termin für den gemeinsamen Umstieg auf die GND soll im Oktober 2011 vereinbart werden.

Weitere Einzelheiten können der Tischvorlage entnommen werden.

## 3.2 Spezielle Themen der Sacherschließung

### 3.2.1 Übergangsregeln

#### Übergangsregeln - Allgemeines

Frau Scheven erläutert Allgemeines zu den GND-Übergangsregeln anhand der entsprechenden Tischvorlage. In den Übergangsregeln wurden die Prinzipien der RDA berücksichtigt, es wurde jedoch versucht, der Entscheidung über RDA nicht vorzugreifen. Abweichungen zur RDA gibt es in Bereichen mit großen kulturellen Unterschieden, hier beschreiben die Übergangsregeln bis zur Einführung eines neuen Regelstandards unsere jetzigen kulturellen Gegebenheiten.

Es wurde darauf hingewiesen, dass zu den Übergangsregeln zu Werktiteln der Musik eine Umfrage bis zum 18.3.2011 stattfindet, zu der auch die Meinung der Kolleginnen und Kollegen der EG Sacherschließung gefragt ist.

#### Übergangsregeln - Vorgehensweise

In der DNB werden die Übergangsregeln zu Individualnamen wie Körperschaften, Geografika, und Personen zurzeit mit den RDA abgeglichen. Die endgültigen Textfassungen aller Übergangsregeln sollen auf der gemeinsamen Sitzung der Expertengruppen Formalerschließung, Sacherschließung und Normdaten im Mai (26./27.) abgestimmt werden, der Versand der Unterlagen soll drei Wochen vor der Sitzung erfolgen.

Die RSWK bleiben bis auf weiteres die Grundlage der inhaltlichen Erschließung. Auf eine Neuformulierung aller betroffenen RSWK-Paragrafen soll aus zeitlichen Gründen verzichtet werden, stattdessen sollen Verweise in der Form von: „Bitte beachten Sie die Übergangsregel ...“ (der genaue Wortlaut muss noch abgesprochen werden) eingearbeitet werden, die den Benutzer zu den entsprechenden Übergangsregeln führen. Das Kennzeichnen aller betroffenen RSWK-Paragrafen soll arbeitsteilig durch die Verbundpartner geschehen. Für alle auf diese Weise geänderten RSWK-Paragrafen soll eine 8. Ergänzungslieferung erstellt werden, die nach Genehmigung durch den Standardisierungsausschuss vor Produktivschaltung der GND, wenn möglich bereits Ende dieses Jahres, erscheinen soll. Sie tritt mit Produktivschaltung der GND in Kraft.

#### RSWK-spezifische Übergangsregeln – Inhalte

Frau Scheven erläutert eine Tabelle (Tischvorlage), in der die Übergangsregeln, die zugehörigen RSWK-Paragrafen und die jeweiligen Beschlüsse aufgeführt sind. Noch zu regelnde Sachverhalte sind grün dargestellt.

##### Mehrgliedrigkeit

Mehrgliedrigkeit soll in der GND weitestgehend aufgelöst werden. Echte Relationen sollen aus den bisherigen Ansatzungs- und Verweisungsformen herausgelöst und in den 5XX-Feldern verknüpft werden. Die DNB möchte deshalb die beschlossenen Übergangsregeln zur Mehrgliedrigkeit auf alle

mehrgliedrigen Ansetzungen (gio, giw, Mundarten u. Ä. ), auf alle 830 p-Verweisungen (Werktitel) und auf alle 830 c+k-Verweisungen (Körperschaften, Veranstaltungen, Schriftdenkmäler) erweitern. Die Angabe von Zusätzen in den Feldern 150/450 \$x bleibt von diesem Vorschlag unberührt.

Die Anwesenden diskutieren diesen Vorschlag kontrovers. Frau Braune-Egloff äußert Bedenken, weil die herausgelösten Relationen als wichtige Suchterme für das Retrieval in den Aleph-Systemen nicht mehr zur Verfügung stehen würden, z. B. Orte sind suchrelevant und könnten nicht durch Codes selektioniert werden. Sie plädiert daher für die bisherige Praxis mit der Dopplung von Informationen in den 4XX-Feldern.

Frau Scheven entgegnet, dass zurzeit viele unterschiedliche Elemente in den 830er Feldern abgelegt werden. Mit den gemeinsamen Regeln für Formal- und Sacherschließung und der Aufhebung der Ortsbindung wird es keine Drehverweisungen mehr geben, ebenso wird die Synonymieverweisung unter dem Sitz bei Körperschaften entfallen. Der Sitz soll in den 5XX-Feldern als spezielle Relation angegeben werden. Das betrifft ca. 150.000 Datensätze. Die vorgeschlagenen Erweiterungen, z. B. für Bauwerke, betreffen wesentlich kleinere Datenbestände.

Einig sind sich die Anwesenden, dass bestimmte 5XX-Relationen für das Retrieval in den Online-Katalogen unabdingbar sind, besonders weil Ortsangaben nicht grundsätzlich als Zusatz im Unterfeld \$g abgelegt werden sollen.

Frau Braune-Egloff befürchtet, dass das Einbinden aller 5XX-Relationen zu großem Ballast führen könnte. Zu prüfen ist, ob in den Aleph-Systemen Indices auch anders aufgebaut werden könnten.

Frau Prof. Wiesenmüller schlägt zwei mögliche Vorgehensweisen vor:

1. Ein automatisches Verfahren könnte suchrelevante Terme aus den 5XX-Feldern in den 4XX-Feldern doppeln.
2. Die überregionale GND bleibt sauber und die Aleph-Verbünde überlegen sich ein eigenes Verfahren.

Frau Pfeifer erinnert daran, dass im Rahmen der Online-Kommunikation Datenbanken kongruent, also 1:1, gehalten werden sollten. Mehrgliedrige Verweisungen, die nicht über die Schnittstelle gehen, regional zu halten, wäre deshalb problematisch. Frau Prof. Wiesenmüller schlägt unter diesen Umständen die maschinelle Doppelung der Angaben als Übergangslösung vor.

Die Anwesenden tauschen sich über Suchmöglichkeiten in den Aleph-Verbänden aus. Neben den eigentlichen OPACs kommt verstärkt Suchmaschinentechnologie zum Einsatz, wodurch die diskutierten Schwierigkeiten in den Aleph-Systemen u. U. relativiert werden können. Im BVB bekommt der Benutzer den Zugang über das Gateway Bayern auf das Aleph-System. Für den KOBV wird die Online-Katalog-Software Primo gerade für den Zugang eingerichtet, deren Retrieval-Möglichkeiten allerdings z. T. noch unklar sind. Im OBV befindet sich Primo bereits im Einsatz.

Für Herrn Dr. Stumpf steht fest, dass sich diese Systeme weiterentwickeln werden. Nach seiner Meinung ist der Benutzer mittlerweile den Umgang mit großen, auch ballasthaltigen Treffermengen gewohnt, die er dann mit Hilfe von Drill-Down-Menüs auf die wesentlichen Inhalte einschränken kann. Demnach könnte das erste Suchergebnis durchaus alle Treffer einschließlich derer zu den 5XX-Feldern enthalten. Außerdem müssen Bauwerke generell mit ihren Ortsangaben suchbar sein. Die verbesserte Strukturierung der Normdaten in der GND soll für intelligentere Sucheinstiege im Online-Katalog genutzt werden.

Es wird vorgeschlagen die Abstimmung bis zur Klärung der technischen Möglichkeiten zu verschieben, jedoch lässt der enge Zeitplan des GND-Projekts keinen Aufschub zu.

Um bisherige Inhalte und Sucheinstiege durch Synonymieverweisungen zu sichern, schlägt Frau Scheven vor, auch in der Migrationsstufe 2 die Informationen der 830-Verweisungen in den 4XX-Feldern zu erhalten. Weitere Datenanalysen sollen zur endgültigen Klärung dieser Frage beitragen.

Zu den Abstimmungen liegt eine Tischvorlage vor. An den Abstimmungen nehmen alle Bibliotheksverbände außer OBV (der jedoch durch die anderen Aleph-Verbände vertreten wurde), SBB und

ekz teil. Als allgemeiner Vorbehalt zu den Regeln wird von den Anwesenden angemerkt, dass die Eindeutigkeit der Ansetzung immer gegeben sein muss.

#### ÜR RSWK MA1.2.2a: Übergangsregel mehrgliedrige geografische Schlagwörter

Erweiterung der Übergangsregel RSWK für mehrgliedrige Ansetzungen (ÜR RSWK MA1.2.2.)

Beschluss war: Die mehrgliedrige, ortsgebundene Ansetzung wird aufgegeben (analog zur Aufgabe der Ortsbindung bei Körperschaften, vgl. TOP am 18.11.2009). Der Standort wird in der 551 als eine spezielle Relation angegeben.

Antrag: Der getroffene Beschluss für Bauwerke, Großplastiken, Grab- und Denkmäler, Sammlungen wird für kleinräumige Geografika (Entitätencode gio), Wege, Strecken, Linien (Entitätencode giw) und Sprachen (Mundarten) erweitert.

Alle anwesenden Verbände sprechen sich einstimmig für den Antrag (10 Stimmen bzw. 11, wenn das Votum der Aleph-Verbände auch für den nichtanwesenden OBV gezählt wird) aus.

#### ÜR RSWK MA1.2.2.1: Übergangsregel für Bauwerke, Großplastiken, Grab- und Denkmäler, kleinräumige Geografika, Wege, Strecken, Linien

Zwei Optionen stehen zur Auswahl:

Option 1: Die Ansetzung erfolgt unter dem Namen; der Standort wird als spezielle Relation angegeben. Wenn der Name alleine homonym zu einer anderen Ansetzung ist, wird der relationierte Ort als Homonymenzusatz hinzugefügt.

Option 2: Die Ansetzung erfolgt unter dem Namen; der Standort wird als spezielle Relation angegeben und grundsätzlich als Zusatz in der Ansetzungsform erfasst.

Die Verbände BSB, BVB, GBV, HBZ, HeBIS, KOBV, SNB, SWB, ZIKG und der OBV (in Abwesenheit, vertreten durch die restlichen Aleph-Verbände) sprechen sich für Option 2 (10 Stimmen) aus, die DNB enthält sich der Stimme.

Ergänzender Antrag: Option 2 gilt gleichermaßen für Synonyme

Die Verbände BSB, BVB, GBV, HBZ, HeBIS, KOBV, SNB, SWB, ZIKG und der OBV (in Abwesenheit, vertreten durch die restlichen Aleph-Verbände) sprechen sich für den Antrag (10 Stimmen) aus, die DNB enthält sich der Stimme.

#### Ländercode

In der Tischvorlage zur Übergangsregel Ländercode (ISO 3166) wird auf den ersten zwei Seiten der Ist-Stand in den drei Normdateien SWD, PND und GKD dokumentiert, der kurz erläutert wird. Der Vorschlag zur Übergangsregel auf Seite drei der Tischvorlage wird mit einer Ergänzung wie folgt angenommen:

1. Abgeschlossene Körperschaften, Kongresse, Gebietskörperschaften und verstorbene Personen bekommen, wenn möglich, den zeitlich zutreffenden Code und den aktuellen Code.
2. Aktuell existierende Geografika bekommen nur den jeweils aktuell zutreffenden Code.
3. Noch bestehende Körperschaften und Kongressfolgen und noch lebende Personen bekommen ebenfalls nur den aktuellen Code. Dabei wird bis auf die Satzart Tp (Personen) auch ISO 3166-2 in der bisherigen Auswahl (Deutschland, Österreich, die Schweiz, Tibet und Südtirol) benutzt (Änderung für die GKD).

Einzelheiten regelt der Leitfaden zur Vergabe des Ländercodes. Dieses Dokument muss vollständig überarbeitet und auch für die bisherigen GKD-Anwender als Grundlage verankert werden.

Wichtige Aspekte aus der Diskussion zu diesem Thema:

Der Vorschlag wird bezüglich der Vergabe von Ländercodes bei Personen diskutiert. In der PND sind zurzeit nur die Ländercodes für Staaten (d. h. ISO 3166-1) ohne ihre Gliedstaaten (ISO 3166-2) zugelassen, im Gegensatz zur SWD wird ISO 3166-2 nicht verwendet. Von Frau Büning wurde der Wunsch geäußert, bei Personen den Ländercode auch auf Gliedstaatenebene anzugeben. Das ist problematisch, da eine Person häufig ihren Ort einschließlich des Gliedstaats wechselt. Es ging ihr eher darum, Personendatensätze, die in der Landesbibliografie verwendet werden, zu kennzeichnen. Das sollte besser über ein Nutzungskennzeichen erfolgen. Das hbz wird die Anforderung nochmal schriftlich formulieren [*am 18.3.2011 per E-Mail an Frau Scheven und Frau Pfeifer erfolgt*].

Frau Prof. Wiesenmüller berichtet von einem neuentwickelten Tool im Online-Katalog der Universität Heidelberg (HEIDI), welches einen Drill-down auf der Basis der Ländercodes in den Normdatensätzen ermöglicht; dabei werden auch die Hierarchien der aktuellen Ländercodes ausgewertet. Sie möchte für diese und ähnliche Anwendungen der Benutzerführung die Hierarchien durch eine möglichst vollständige Verwendung der aktuellen Codes erhalten; sie beantragt deshalb die Doppelcodierung mit zeitlich zutreffenden und aktuellen Codes bei abgeschlossenen Körperschaften, Kongressen, Gebietskörperschaften und verstorbenen Personen. Der aktuelle Code soll dabei immer vergeben werden, der zeitlich zutreffende Code kann dazu treten. Die anwesenden Verbände stimmen dem Antrag einstimmig (11 Stimmen; der OBV wird durch die anwesenden Aleph-Verbände mitvertreten) zu.

Mehrfachkodierungen sind grundsätzlich möglich, bis zu vier Codes sind wie bisher zulässig. Zur Zeit sieht die Planung so aus, dass bei der Zusammenführung von Normdatensätzen aus GKD und SWD alle unterschiedlichen Codes der einzelnen Datensätze in den neuen Datensatz übernommen werden. Einzelheiten wie der Umgang mit mehr als vier Codes müssen noch geklärt werden.

Offene Fragen zum Ländercode:

Die Anwesenden überlegen, wie man in der GND mit „diplomatischen Vertretungen“ umgehen will. In der SWD werden sie doppelt codiert, sie erhalten den Code des entsendenden Landes und des Gastlandes. Nach Stichproben vergibt die GKD nur einen Code, den Ländercode des entsendenden Landes. Eine einheitliche Regelung wird gewünscht, möglichst entsprechend der jetzigen SWD-Regelung. Die Anwendung des Ländercodes innerhalb der GKD muss geklärt werden.

Die GKD vergibt außerdem nur Codes auf Staaten-Ebene. Für den Teilbestand „s“ ohne Teilbestand „f“ soll an der jetzigen SWD-Regelung festgehalten werden.

Die GKD-Codes DXDE und AAAT (Österreich-Ungarn) benötigen die SWD-Redakteure nicht.

### 3.2.2. Splits bei Gebietskörperschaften

Im Anschluss an die letzte Expertengruppen-Sitzung wurde eine Umfrage zu Splits bei Gebietskörperschaften durchgeführt; dabei traten neue Überlegungen auf, die in einem Papier von Frau Prof. Wiesenmüller zusammengefasst sind (vgl. die entsprechende Unterlage zur Tagesordnung). Formuliert wurden dabei zwei grundsätzliche Optionen: Bei der Option 1, die sich an der Praxis der LCSH orientiert, würde bei Gebietskörperschaften im Bereich der Sacherschließung (anders als in der Formalschließung) weiterhin der jeweils jüngste Name verwendet; bei der Option 2 gäbe es eine konsequente Splitting-Lösung auch für Gebietskörperschaften. Die Verbände sprechen sich ausnahmslos für die Option 1 aus, d. h. die Beschlagwortung erfolgt mit dem neuesten Namen und entspricht damit eher der heutigen Praxis, auch wenn in der GND zukünftig gesplittete Datensätze vorhanden sind. Welcher Datensatz für die Sacherschließung zu verwenden ist, soll durch einen Verwendungshinweis im Datensatz transportiert werden. Ändert sich der Name der Gebietskörperschaft, müssen diese Hinweise aktualisiert werden.

Die unterschiedliche Behandlung von Gebietskörperschaften und normalen Körperschaften wird diskutiert. Da Datenanalysen fehlen, wissen die Anwesenden nicht genau welche Datensätze es in

welchem Umfang betrifft. Da Gebiete laut Frau Pfeifer in der Regel keine Akteure sein sollten, ist eine unterschiedliche Behandlung gerechtfertigt. Auch Frau Prof. Wiesenmüller sieht einen unterschiedlichen Sachverhalt gegeben und hält somit eine unterschiedliche Behandlung für plausibel; dies entspricht auch der angloamerikanischen Praxis.

Für die restlichen Körperschaften muss später, vor Match & Merge, überlegt werden, wie sie für die Sacherschließung umgesetzt werden und wie das Retrieval gestaltet wird. Das Match-Merge-Verfahren zwischen SWD und dem neusten GKD-Satz kann in der DNB im Moment noch nicht konzipiert werden. Betroffene Gebietskörperschaften enthalten noch keine Benutzungshinweise. Die DNB versucht, eine Liste der betroffenen Gebietskörperschaften zu erstellen.

Abstimmung: Die anwesenden Verbände stimmen einstimmig (11 Stimmen, der OBV wird durch die Aleph-Verbände vertreten) für Option 1.

### 3.2.3. Mehrgliedrige Oberbegriffe

Im Vorfeld dieser Sitzung erfolgte im Wiki eine Umfrage zum Papier "Vorschlag zur Umsetzung der MOs in der GND" (<https://wiki.d-nb.de/x/6QmKAg>); der vorgeschlagenen Vorgehensweise wurde grundsätzlich zugestimmt. In einer Tischvorlage sind alle Kommentare der Verbände einschließlich der Antworten der DNB dargestellt.

Der Vorschlag sieht die Auflösung aller mehrgliedrigen Oberbegriffe und die Relationierung ihrer einzelnen Bestandteile in den 5XX-Feldern vor. Nicht gefundene Ansetzungsformen bleiben als Text in 5XX stehen und können händisch aufgearbeitet werden. Das Unterfeld „\$Y“ der 5XX-Felder für die Kennzeichnung der ehemaligen MOs entfällt. Die Verbände können sich bei Bedarf aus bestimmten Codierungen eine Art MO zusammenstellen. Das Geografikum des MOs wird nicht, wie ursprünglich geplant, durch den Ländercode ausgedrückt, sondern auch in den 5XX-Feldern relationiert mit einem neuen Code „geoa“ versehen.

### 3.2.4. Umsetzung von Mundarten

Als Tischvorlage dient ein Vorschlagspapier zur Datenbereinigung von mehrgliedrigen Oberbegriffen bei Mundarten und Sprachen. Der Vorschlag sieht die konsequente Löschung aller mehrgliedrigen Oberbegriffe vor. Die Löschung in den 71 betroffenen Datensätzen soll manuell vor der Migration vorgenommen werden, bei Bedarf wird der MO in echte Oberbegriffe umgewandelt.

Mehrgliedrige Oberbegriffe im SWD-Feld 850, das betrifft 140 Datensätze, werden bei der Migration durch das Fehlerprotokoll abgefangen und müssen dann manuell korrigiert werden.

### 3.2.5. Liste der \$4-Codierungen

Im Vorfeld dieser Sitzung fand im Wiki auch eine Umfrage zur Codeliste der Relationstypen statt (<https://wiki.d-nb.de/x/qQaKAg>). Eine Tischvorlage mit Änderungsvorschlägen zur Codeliste wird ausgeteilt.

Frau Scheven erläutert, wie die Codeliste entstanden ist. Zuerst wurden die dreistelligen Codes der MARC-Codeliste in Hinblick auf die GND-Bedürfnisse untersucht. Die MARC-Codes bedienen aber andere Funktionen und viele GND-Anwendungen fehlen. Unklar waren und sind auch die Auswirkung der RDA-Einführung auf die Codes. Deshalb wurden für den deutschen Sprachraum und den deutschen Gebrauch neue sprechende Abkürzungen definiert.

Herr Heuvelmann erläutert, dass für die Festlegung der MARC-Codes kein Komitee zuständig ist. Ansprechpartnerin für die GND-Codeliste, die (auf dem Stand Ende 2010) auch schon in englischer Sprache vorliegt, ist Sally McCallum (Chief of the Network Development and MARC Standards Office of the Library of Congress). Grundsätzlich ist man dort gegenüber neuen Entwicklungen sehr aufgeschlossen. Die RDA sind im Bereich der Sacherschließung noch sehr lückenhaft; im amerika-



nischen Raum werden dazu Diskussionen erwartet. Aber auch die bestehenden Anhänge I, J und K, die die Relationen auflisten, sind noch in Bewegung, so dass hier umfangreiche Ergänzungen bei den MARC-Codes zu erwarten sind. Mit der GND und ihren Relationen ist man im deutschsprachigen Raum der Entwicklung voraus. Herr Heuvelmann erwartet deshalb keine großen Überschneidungen mit anderen Anwendungen, rechnet aber andererseits damit, dass die Vergabe von Codierungen zur Abbildung von Relationen in die internationale Regelwerks- und Formatentwicklung einzubinden ist.

In Absprache mit Sally McCallum werden von offizieller Seite dem deutschsprachigen Raum offizielle MARC-Codes zugewiesen werden. Die vierstelligen deutschsprachigen GND-Codes sind demnach noch keine internationalen Codes. Im MARC-Austauschformat, einschließlich Online-Schnittstelle, dürfen aber nur internationale Codes ausgetauscht werden. Die DNB schlägt deshalb ein Mapping 1:1 vor dem Austausch und ein umgekehrtes Mapping für den Import vor.

Frau Kunz spricht sich für die ausschließliche Verwendung der internationalen MARC-Codes und gegen eine Konversion aus. Sie findet die vorgelegte Liste mnemotechnisch auch nicht durchgängig logisch.

Als Austausch-Alternative der \$4-Codierungen wird die anwenderspezifische Auslieferung im MARC-Unterfeld „\$94:“ vorgeschlagen. Die EG Datenformate hat auf ihrer Sitzung am 17.02.2011 für das MARC-Austauschformat pauschal \$9 definiert.

Herr Heuvelmann schlägt vor, die Codeliste in vier Großgruppen analog der RDA aufzuteilen:

- Name – Name (Appendix K)
- Resource – Resource (Appendix J)
- Name – Resource (Appendix I)
- Other (Appendix L und ggf. weitere)

Die differenzierte Ausweisung analog der RDA wird diskutiert. Folgendes Meinungsbild wird zu diesem Vorschlag festgehalten. Sieben Verbände (BSB, BVB, HBZ, HeBIS, KOBV, SNB, SWB) sprechen sich für die unstrukturierte Darstellung aus, drei Verbände enthalten sich (ZIRK, DNB, OBV) und der GBV ist für die differenzierte Ausweisung.

Alle neu vorgeschlagenen Codes der Tischvorlage werden akzeptiert und sollen in die Liste übernommen werden. Entsprechend dem MO-Papier wird der Code „re1t“ nicht eingeführt, sondern „them“ verwendet.

Folgende Vorgehensweise wird festgehalten:

1. Endgültige Abklärung (Codes und ihre Darstellung in der Liste, Unterfeld im Austauschformat) in allen Expertengruppen bis Ende April
2. Übersetzung der Liste ins Englische mit der gewünschten Zuordnung nach Großgruppen und Versenden der Liste vor der ALA an Sally McCallum
3. Absprachen von Herrn Heuvelmann mit Sally McCallum beim ALA-Treffen „Annual Conference and Exhibit“ vom 23.-28. Juni 2011 in New Orleans.

### 3.2.6. Umsetzung von Hinweissätzen

In den verschiedenen Migrationsstufen werden alle SWD-Hinweissätze standardmäßig 1:1 in GND-Hinweissätze umgewandelt. Die enthaltenen IDNs bleiben dabei erhalten. Alle Hinweissätze werden in Pica in dem Feld „Satzart“ an Position 4 mit einem „e“ gekennzeichnet.

Gemäß ÜR RSWK MA1.2.2 werden bestimmte Typen von Hinweissätzen (Sammlungen, Grab- und Baudenkmäler ...) zu eigenen Datensätzen. Wenn mehrere Hinweissätze zu einer Entität gehören, müssen diese mit einem Match-and-Merge-Verfahren zusammengeführt werden. Dabei werden alle Datensätze mit der gleichen Kombination zu verknüpfender Deskriptoren (606er Felder) so zu einem Datensatz zusammengefasst, dass ein 1XX-Feld und mehrere 4XX-Felder entstehen. Näheres anhand von Beispielen kann der Tischvorlage entnommen werden.

Dieses Match-and-Merge-Verfahren erfolgt wie alle anderen Match-and-Merge-Verfahren erst nach der Produktionsaufnahme der GND.

### 3.2.7. Match-and-Merge

Der Schritt „Match-and-Merge“ soll erst nach Produktivschaltung der GND geschehen. Die in der Tischvorlage aufgelisteten Mengenabschätzungen werden kurz besprochen. Detaillierte Anweisungen für das Verfahren werden erst zu einem späteren Zeitpunkt vorgestellt, da die Auswertung der Datenanalysen noch nicht fertig ist. Den größten Bereich bilden die Körperschaften. Die Zusammenführung von GKD- und SWD-Sätzen soll über das eindeutige SWD-Feld 028 (GKD-Nummer) erfolgen. Wenn die Produktivschaltung der GND im ersten Quartal 2012 erfolgt, wird das Match-and-Merge-Verfahren voraussichtlich im zweiten Quartal 2012 umgesetzt.

### 3.2.8. Geografikum nach Formschlagwörtern (z. B. Kongress)

In Schlagwortfolgen der Titeldaten werden Geografika nach Formschlagwörtern zurzeit als Freitext erfasst. Eine Ausnahme bildet derzeit noch der SWB, der für solche Geografika, aber auch für Zeit- und Formschlagwörter regionale Schlagwörter anlegt, mit denen verknüpft wird; diese Praxis wird aufgrund des Übergangs in die GND jedoch in Frage gestellt. Die derzeitige Situation zeigt, dass die bisherige Praxis suboptimal ist: Aufgrund der GND-Übergangsregeln werden sich viele Ansetzungen von Geografika ändern; dies kann in den Freitextangaben der Schlagwortfolgen in den Titeldaten nicht sofort nachvollzogen werden wie bei den Normdatenverknüpfungen. Deshalb schlägt DNB für Geografika nach einem Formschlagwort die Verknüpfung mit einem Normdatensatz vor. Die in diesem Feld befindlichen Zeitangaben sollen als ein neues Element der Schlagwortfolge, eingeleitet mit dem Indikator „z“, abgelegt werden.

Der Wunsch nach einer Normdatenverknüpfung besteht auch für Form- und Zeitschlagwörter; auch hier ist die derzeitige Situation nicht befriedigend (insbesondere wegen der nicht suchbaren Verweisungsformen bei den Formschlagwörtern). Allerdings wird gewünscht, dass eine Kennzeichnung der jeweiligen Funktion in der Schlagwortfolge vorhanden ist, um nicht zutreffende Ergebnisse beim Retrieval auszuschließen (beispielsweise soll eine inhaltliche Suche zu einem Geografikum nicht durch Treffer, in denen das Geografikum als Kongressort angegeben wird, verfälscht werden). Als Möglichkeit wird ein Rollenoperator vorgeschlagen.

Die DNB beabsichtigt, die Umsetzung dieses Beschlusses in den Katalogdaten nach Produktivschalten der GND Ende des 1. oder Anfang des 2. Quartals 2012 vorzunehmen. Frau Büning fragt, ob die Verbände ihre Titel ebenfalls korrigieren müssten und ob das für den Datenaustausch wichtig wäre. Es handelt sich um große Titelmengen, in denen korrigiert werden müsste. Das Herstellen der Verknüpfung zu den entsprechenden Normdatensätzen wäre sehr aufwändig. Auch der BVB hält eine Korrektur der Altdaten für zu aufwändig.

Die technischen Möglichkeiten zur Kennzeichnung von unterschiedlichen Rollen der Schlagwörter in der Schlagwortfolge müssen geklärt werden, ebenso die Problematik der Altdaten. Solange wird der Antrag zurückgestellt.

Grundsätzlich besteht jedoch Interesse daran, sich nach der Einführung der GND wieder mit dem Thema zu beschäftigen. Im Wiki der EG SE soll ein Diskussionsforum eingerichtet werden, um Perspektiven dafür zu entwickeln.

## 4. Automatische Erschließung für Netzpublikationen

Vor dem Hintergrund des erweiterten Sammelauftrags (Sammlung und Erschließung von Netzpublikationen) steht die DNB vor der Frage, wie sie bei gleichbleibenden oder schrumpfenden Ressourcen eine stetig steigende Zahl an Publikationen sammeln und bearbeiten kann. Im DNB-internen Petrus-Projekt werden seit Anfang 2009 automatische Verfahren entwickelt, die auf die Erschlie-

Bung von Netzpublikationen ausgerichtet sind. Frau Junger gibt einen Überblick zum Stand des Projekts (Präsentation siehe Wiki).

Das Projekt gliedert sich in vier Anwendungsszenarien, die als Teilprojekte realisiert werden.

- Sz1: Automatische Vergabe der DNB-Sachgruppen
- Sz2: Normdatenrelationierung
- Sz3: Parallelausgaben
- Sz4: Automatische Beschlagwortung

Im Szenario 2 wurde unter Einsatz der PICA-Match-und-Merge-Software ein Verfahren zur automatischen Verknüpfung von Personennamen in Titeldaten zur Personennamendatei (PND) entwickelt. Am 1. März 2011 wurde die automatische PND-Verknüpfung (Stufe 1) implementiert. Personennamen, die als Textphrase vorliegen, werden mit den Ansetzungsformen der PND-Einträge abgeglichen, wobei in Stufe 1 nur Tn-Sätze berücksichtigt werden. Wenn kein Tn-Satz passt, wird automatisch ein Tn-Satz angelegt und mit Katalogisierungslevel 7 gekennzeichnet.

Im Rahmen des Szenarios 3 wurde unter Einsatz der PICA-Match-und-Merge-Software ein automatisches Verfahren zur Erkennung von parallelen Print- und Onlineausgaben entwickelt. Ziele sind die automatische Verknüpfung von parallelen Print- und Onlineausgaben und die maschinelle Übernahme bereits vorhandener Inhaltserschließungsdaten und Normdatenverknüpfungen in die Titelaufnahmen der entsprechenden parallelen Ausgaben. Seit 1. März ist auch dieses Verfahren im Einsatz.

Unklar ist, ob und wie automatisch erzeugte Daten in den Datentausch gehen *[Anmerkung: Automatisch transferierte Sacherschließungsdaten, Verknüpfungen zwischen Parallelausgaben sowie automatisch erzeugte Verknüpfungen zur PND werden ohne besondere Kennzeichnung ausgeliefert.]*

Die Szenarien 1 und 4, in die die Abteilung Inhaltserschließung eingebunden ist, befinden sich noch in der Entwicklungsphase. Grundlage ist eine Software der Firma Averbis GmbH aus Freiburg zur automatischen Erschließung.

Im Szenario 1 (automatische Sachgruppenvergabe) soll ein automatisches Klassifizierungsverfahren zum Einsatz kommen, das mit Hilfe bereits erschlossener Dokumente trainiert wird. Ziel ist eine automatische Vergabe von DDC-Sachgruppen für möglichst alle Fächer. Trainiert wurde das System einerseits mit Volltexten, andererseits auch mit digitalisierten Inhaltverzeichnis. Das Qualitätsziel ist ein F-Measure (harmonisches Mittel aus Recall und Precision) von 0,8, verglichen mit einer Erschließung auf intellektueller Basis. Probleme ergeben sich in den Sachgebieten, in denen es nicht genug Trainingsmaterial gibt.

Ziel von Szenario 4 ist die automatische Anreicherung der Titeldaten deutschsprachiger Online-Publikationen mit SWD-Schlagwörtern und freiem Vokabular. Tests erfolgten bisher anhand eingliedriger Sachschlagwörter der SWD mit ihren verzeichneten Relationen. Die Ergebnisse werden intellektuell bewertet. Statistische Methoden bilden die Grundlage zur Ermittlung inhaltsrelevanter Deskriptoren, wegen der schwer zu analysierenden deutschen Sprache können linguistische Verarbeitungsschritte hinzutreten.

Zur Unterstützung der Geschäftsprozesse entsteht auf diese Weise Schritt für Schritt ein modular zusammengestelltes und nachregelbares Erschließungssystem, das zwischen verschiedenen Medientypen und deren Erschließungsanforderungen differenzieren kann und dessen Ablauf über definierte Qualitätskriterien gesteuert wird. Laut Frau Junger soll zukünftig immer mehr automatisch erschlossen werden. Das Projekt PETRUS endet im Dezember 2011.

Herr Dr. Eberhardt regt an, eine Feedback-Funktion in den Online-Katalog einzubauen. Herr Holbach schlägt vor, zu untersuchen, ab wann die weitreichende Anreicherung der Katalogdaten mit inhaltstragenden Begriffen, z. B. auch freien Schlagwörtern, schädlich für das Retrieval sein könnte. Frau Junger erwidert, dass der Treffersortierung nach Relevanz (Relevanz-Ranking) in Bibliothekskatalogen immer größere Bedeutung zukomme. Die Anwesenden sind an den Testergebnissen interessiert.

## 5. Verschiedenes

### 5.1 Rangfolge der Nachschlagewerke

Probleme mit dem elektronischen Tool zur „Liste der fachlichen Nachschlagewerke“ können in Abwesenheit von Frau Dr. Sandner nicht besprochen werden. Eine Mail von Frau Dr. Sandner zu diesem Thema soll weitergeleitet werden. *[zwischenzeitlich erfolgt]*

### 5.2 RSWK-Änderungen

Der Vorschlag zur Änderung RSWK § 701,1c, Abs. 2 von Herrn Dr. Horstkotte und Frau Dr. Nadj-Guttandin soll per Umlaufverfahren abgestimmt werden.

### 5.3 Umfragebogen

An die DNB wurde ein Fragebogen über eine Umfrage zu Konkordanzen von RVK mit DDC geschickt mit der Bitte, diesen an die EG SE weiterzuleiten. Das soll erfolgen und wurde auf der Sitzung angekündigt. *[zwischenzeitlich erfolgt]*

### 5.4. Retrieval-Möglichkeiten

Frau Prof. Wiesenmüller schlägt vor, Ideen zum Retrieval zentral zu sammeln und in den § 20 „Gestaltung der Schlagwort-Recherche im Online-Katalog“ RSWK aufzunehmen. Im Wiki soll ein Bereich dafür eingerichtet werden.

### 5.5 Sachgruppen in MAB2-Titel

Herr Heuvelmann erläutert an Hand einer Präsentation "MAB-Feldbelegung für Sachgruppen der DNB und Sachgruppen der ZDB" die im Sommer 2010 zwischen der ZDB und der DNB vereinbarte und gültige Praxis. In MAB Feld 700 "Notation eines Klassifikationssystems" sind der Indikator "d" für die Angaben nach der SWD-Systematik (mit Bezug zu den Schlagwortfolgen nach RSWK), und der Indikator "z" für die Angaben nach der ZDB-Systematik belegt. Der Indikator Blank für eine „Systematik der katalogisierenden Institution“ wird seit jeher mit den Sachgruppen der Deutschen Nationalbibliografie belegt, mit allen vorhandenen zeitlichen Stufungen (bis 1981, ab 1982, und aktuell mit den DDC-basierten Sachgruppen). Zur Unterscheidung, ob eine DDC-basierte Sachgruppe von der DNB, oder aber von der ZDB erfasst wurde, werden seit August 2010 DDC-basierte Sachgruppen der DNB und DDC-basierte Sachgruppen der ZDB nach dem Muster „700\_|Notation#DNB“ und „700\_|Notation#ZDB“ voneinander unterschieden und in MAB transportiert. Mit Hilfe des in MAB als Strukturelement reservierten Teilfeldtrennzeichens (hier dargestellt als "#") und dem Institutionenkennzeichen wird dabei für eine eindeutige Kennzeichnung des Erschließungskontextes gesorgt.

Die Lieferung in MAB 700\_Blank führt laut Herrn Holbach zu Problemen in den Verbänden, da diese dort Notationen von verbundinternen Systematiken führen und u.a. Validierungen und Überschreibsperrern mit dem Feld verknüpft sind. Er regt an, die DDC-Sachgruppen der DNB und ZDB in den im MAB-Format eigens definierten Feldern MAB 700d („Systematik der Deutschen Nationalbibliographie“) und MAB700z („ZDB-Systematik“) auszuliefern. Zwar sind in älteren Titelaufnahmen der DNB dort früher Notationen der SWD bzw. der alten ZDB-Systematik abgelegt worden, aber man könnte dort die Sachgruppen, die ja ein auch klassifikatorisches Element darstellen, von den bereits vorhandenen Notationen leicht durch ein Suffix unterscheiden, sofern man die alten Notationen nicht löschen will/kann.

Herr Heuvelmann argumentiert gegen eine Änderung der jetzigen Praxis. Sie ist zwischen den Partnern ZDB und DNB abgestimmt, für die aktuelle Sitzung ist dies noch einmal mündlich bekräftigt worden. Es handelt sich zudem um eine Lösung für die DNB- und ZDB-Lieferungen im Austauschformat MAB, d.h., dass es Beziehern frei steht, beim Import in das jeweilige System mit Hilfe der eindeutigen Kennzeichnungen die Daten den internen System-Vorgaben gemäß (wie Validierungen, Indexierungen, etc.) zu strukturieren. Dies kann nicht Aufgabe einer liefernden Institution sein. Das Modell, in 700z die Angaben nach der bisherigen ZDB-Systematik einerseits, und die Angaben nach den DDC-basierten Sachgruppen, vergeben von der ZDB, andererseits zu liefern, oder dann zu halten, birgt das erhebliche Risiko, dass die Kennzeichnung per Suffix ignoriert und so die unterschiedlichen Bedeutungen derselben Ziffernfolge verwechselt werden. Allgemein, so Herr Heuvelmann, ist die Diskussion dieser Frage nicht unbedingt in der Expertengruppe Sacherschließung richtig angesiedelt; sie gehört seiner Meinung nach eher in den Kontext der Expertengruppe Datenformate.

Die DNB sagt eine weitergehende Klärung zu. Die Folien zur Präsentation stehen im Wiki zur Verfügung.