

Die Planungen zur Implementierung und Anwendung der GND im BVB



Werner Holbach

4. Leipziger Kongress für Information und Bibliothek
Normdatenanwendertreffen 15. März 2010

Agenda

- Normdaten in der gemeinsamen Verbunddatenbank des BVB und KOBV
- Die Frage des zukünftigen Datenformats
- Notwendige Anpassungen mit Einführung der GND
- Beteiligung an der AG GND
- Datenbereinigungen im Vorfeld
- Lieferung regionaler Normdatensätze an die überregionalen Normdateien

Normdaten in der gemeinsamen Verbunddatenbank des BVB und KOBV

- eigene „libraries“ im Verbundsystem (ALEPH 500 von ExLibris)
- **BayPND**
 - komplette (Ü)PND plus ca. 2,2 Mio. regionale Datensätze;
insgesamt: ca. 5,8 Mio. Datensätze
 - regionale Namensätze (Tn-Sätze) manuell angelegt oder bei Einspeicherungen maschinell erzeugt
- **BayGKD**
 - komplette (Ü)GKD plus ca. 110.000 regionale Datensätze;
insgesamt: 1,4 Mio. Datensätze
- **BaySWD**
 - enthält komplette (Ü)SWD (ca. 870.000 Datensätze);
keine regionalen Datensätze
- Verknüpfung zu den Titeldaten über **ID-Nummer + Ansetzungsform**

Die Frage des zukünftigen Datenformats

Zukünftiges Intern- und Katalogisierungsformat in der BVB/KOBV-Verbunddatenbank?

Randbedingungen:

- MARC21 ist neues deutsches **Austauschformat**; auch Standard bei **Fremddaten**
- GND wird ab Einführung im Jahr 2011 **ausschließlich im MARC21-Format** ausgeliefert
- **Titeldaten** der DNB und der Bibliotheksverbände können in MARC21 und für (begrenzte Zeit) in MAB2 ausgetauscht werden

Die Frage des zukünftigen Datenformats

Variante 1: Weiterbetreiben der Datenbank auf **MAB2-Basis**

- **Vorteile:** bestehendes System mit ausgereiften Funktionalitäten; Versorgung der Lokalsysteme ebenfalls
- **Nachteile:**
 - Konversion der GND-MARC-Datensätze auf MAB-Datensätze und Aufteilung auf drei getrennte Normdateien (PND, GKD, SWD): große Schwierigkeiten, Datenverluste
 - Online-Kommunikation müsste wieder aufgegeben werden
 - Probleme bei der Realisierung zukünftiger Änderungen bei Regelwerk und Katalogisierungskonventionen
 - Datentausch sehr erschwert

Die Frage des zukünftigen Datenformats

Variante 2: Komplettumstieg auf MARC21 (Version der EG Datenformate = „D-MARC“)

- **Vorteile:**
 - einheitliche Systemarchitektur
 - problemloser Datenaustausch im deutschsprachigen Bereich
 - Realisierung von Verknüpfungen über ID-Nummern, hierarchischen Verknüpfungen bei mehrbändigen Werken und Serien, RSWK-Schlagwortfolgen
- **Nachteile:**
 - deutscher Sonderweg; beim Internationalen Datenaustausch können die „deutschen“ Spezifika nicht getauscht werden
 - redundante Haltung der Schlagwörter (MARC 689 und MARC 600-651)
 - Firma ExLibris ist nicht bereit, diese Variante zu realisieren

Die Frage des zukünftigen Datenformats

Variante 3: Komplettumstieg auf MARC21 (Standard-MARC)

- **Vorteile:**
 - einheitliche Systemarchitektur
 - problemloser Datenaustausch auch mit internationalen Partnern, insbesondere auch für Fremddatennutzung
 - Mitglied der großen Community der ALEPH-MARC-Anwender (Weiterentwicklungen von Aleph, Originalschriftlichkeit, RDA-Umsetzung, Umstieg auf Nachfolgesystem, ...)
 - MARC21 Standard wird (neben anderen) auch vom Nachfolgesystem URM noch unterstützt (zumindest anfangs)
- **Nachteile:**
 - Austauschformat „D-MARC“ erfordert Anpassungen bei Im- und Export
 - RSWK-Folgen können nicht geliefert werden
 - keine einheitliche Systemarchitektur der deutschsprachigen Aleph-Verbünde
 - Realisierung der notwendigen und wünschenswerten Funktionalitäten wäre aufwendig

Die Frage des zukünftigen Datenformats

Variante 4: MAB2-Titel und MARC21-Auth („Hybridlösung“)

- **Vorteile:**
 - „Gestuffer“ Umstieg auf MARC
 - einheitliche Vorgehensweise der Aleph-Verbünde
 - überschaubarer finanzieller Aufwand (insbesondere Anpassung der Schnittstellen und Importroutinen für die GND)
- **Nachteile:**
 - Risiko von bisher nicht absehbaren grundsätzlichen Schwierigkeiten
 - Wie wird GND in den Lokalsystemen realisiert?

Die Frage des zukünftigen Datenformats

- Variante 1 und 2 kommen nicht zum Zug
- **GND wird auf MARC21-Basis realisiert**
- dazu voraussichtlich gemeinsamer Auftrag der deutschsprachigen ALEPH-Verbände an Firma ExLibris

Notwendige Anpassungen mit Einführung der GND

- Realisierung des **Datei-Imports** (inkl. Umlenkungen, Korrekturen, Löschungen etc.)
- Einrichtung des **OAI-Harvesters**
- Einrichtung von Z39.50update-Schnittstelle/**Online-Kommunikationsschnittstelle**
- Lokalsystemfunktionen, bzw. **Versorgungsschnittstellen** zu Lokalsystemen
- Realisierung anwendungsspezifischer **Funktionalitäten** der effektiven Katalogisierung in den Set-ups

Beteiligung an der AG GND

- Mitarbeit von Vertretern des BVB und der BSB an der Arbeit der übergreifenden AG Gemeinsame Normdatei
 - Mitarbeit in den früheren Arbeitsgruppen „**PERSO**“ und „**GKR**“ (Angleichung der Regeln für Formal- und Sacherschließung)
 - Kritische Begleitung (**Stellungnahmen, Änderungsvorschläge, ...**) der Entwürfe der Arbeitsstelle für Standardisierung
 - › **Datenformat**
 - › **Übergangsszenario** für die Implementierung der GND
 - › **Übergangsregeln** für die Katalogisierung (ab GND-Einführung; vor RDA-Einführung)

Datenbereinigungen im Vorfeld: Maschinelle Bereinigungen in der BayPND und BayGKD

- ca. **2,5 Millionen** rein regionale Personennamensätze (manuell angelegt oder durch Einspeicherungen maschinell generiert);
ca. **50.000** regionale Körperschaftssätze
- Ziel: Reduzierung der Bay-PND-Sätze durch Beseitigung von Dubletten

Datenbereinigungen im Vorfeld: Stufe 1, 2 und 3

- Bereinigung falscher **Prüfziffern** („x“ → „X“)
- Herstellung **fehlender Verknüpfungen**:
Ersetzung der **Ansetzungsform** in den Feldern 100ff in den Titeldatensätzen durch die Ansetzungsform aus den Feldern 800 der jeweiligen Normdatensätze
(In ALEPH erfolgt die Verknüpfung über ID-Nummer UND Ansetzungsform!)
- **Löschung** aller Normdatensätze mit einer BV-ID-Nummer, mit denen keine Titeldatensätze verknüpft waren

Datenbereinigungen im Vorfeld: Stufe 4a und 4b

- **Abgleich innerhalb der regionalen Namensätze** und Zusammenführung (mit Umlenkung der Titeldaten) wenn sie ansetzungsgleich waren
(Ursache: Dateneinspeicherungen neuer Verbundbibliotheken)
- **Abgleich der verbliebenen regionalen Namensätze mit (Ü)PND-Namensätzen** und Zusammenführung (mit Umlenkung der Titeldaten), wenn sie ansetzungsgleich sind.

Datenbereinigungen im Vorfeld: Manuelle Bereinigungen

- BayPND und BayGKD: **Fehlerlisten** aus der Maschinellen Bereinigungsaktion wurden abgearbeitet
- BayPND: **regionale Personensätze** (Tp-Sätze) wurden aufgelöst (mit PND abgeglichen und ggf. neu eingebracht oder umgelenkt)
- BayPND: Namensätze aus einer früheren **DMA-Einspielung** wurden auf PND-Sätze umgelenkt
- GKD: Körperschaftsliste der DNB wurde auf **Gebietskörperschaften** überprüft (Vorarbeit für die geplante Zusammenführung von SWD und GKD)
- SWD: seit ca. 3 Jahren werden **Entitäten-Codierungen** eingebracht (Vorarbeit für die geplante Zusammenführung von SWD und GKD)

Lieferung regionaler Normdatensätze an die überregionalen Normdateien

- Ziel: 1:1-Spiegel der überregionalen Normdatei(en) in der Verbunddatenbank
- (Entknüpfung und) **Löschung** der (1,4 Mio.) BayPND-Sätze, die nur mit **genau 1 Titel** verknüpft sind („Eintagsfliegen“)
- **Einspeicherung** der übrigen (ca. 700.000) BayPND-Sätze in die PND (Match-&-Merge)
- Lieferung an **VIAF**; Hinzuziehung der Titelinformationen mit dem Ziel automatischer Umlenkungen und Satzartänderungen (Tn → Tp)
- Abgabe von ca. 50.000 **BayGKD-Sätzen** an die GKD
- Termin: vor Einführung der GND

Vielen Dank!



werner.holbach@bsb-muenchen.de