

Spartenübergreifende Nutzung der GND : am Beispiel des Projekts IN2N

Brigitte Wiechmann, DNB

IN2N - Projekt

IN2N = Institutionenübergreifende Integration von Normdaten

DFG-Projekt, Laufzeit 01.12.2012-28.02.2014

Partner:

Deutsches Filminstitut (DIF), <http://deutsches-filminstitut.de/>

Normdatenbestand: 188.000 Personendatensätze

Deutsche Nationalbibliothek (DNB), www.dnb.de

Die GND umfasst mehr als 11 Mio. Datensätze, davon 3,8 Mio. Personendatensätze

IN2N – Ziel

Ziel des Projekts:

Domänenübergreifende Normdatenpflege im eigenen System

Kooperation soll als Modell für nicht-bibliothekarische Einrichtungen dienen, die Erschließung mithilfe von Normdaten betreiben und von der in der GND bereits geleisteten Arbeit profitieren können

Datenbank des DIF



- www.filmportal.de
- Beispiel: http://filmportal.de/person/mario-adorf_8c9ba38d9d5d4a6d957251f7386471e5

IN2N - Herausforderungen

- Unterschiedliche Regeln, die der Erfassung zugrunde liegen (RDA, Hausregeln, orientiert an EAC-CPF)
- Unterschiedliche Formate und Redaktionssysteme (GND: MARC-XML, DIF: EAC-CPF/XML)
- DIF benötigt nicht den vollen Umfang der GND-Datensätze
- Mitarbeit in der GND erfordert formale Voraussetzungen, die außerhalb des Bibliothekswesen nicht bekannt sind oder nicht angewandt werden

IN2N – Vorgehen

Initiale Einspielung der DIF-Datensätze in die GND mit Match & Merge

Kriterien dazu sind in einem Katalog festgelegt

3 Ergebnisfälle:

- Treffer mit einem vorhandenen GND-Satz = zusätzliche Informationen aus dem DIF-Satz werden in den GND-Satz eingetragen und umgekehrt
- Kein Treffer = neuer GND-Satz wird gebildet, die GND-Nr. wird an DIF gemeldet
- Ein oder mehr als ein möglicher Treffer = Datensatz kommt in die Kandidatenliste und muss intellektuell bearbeitet werden

Initialer Datenabgleich - Matchbedingungen

Elementtyp	Vorgehen
Namen	<ul style="list-style-type: none"> • Identifikation von Namenstypen • Zerlegung in Bestandteile • Zuordnung von GND- und DIF-Bestandteilen • Erstellung kontrollierter Listen für akademische Grade und Adelstitel
Datum	<ul style="list-style-type: none"> • Syntaxanpassung • Umgang mit ungenaueren Angaben
Orte	<ul style="list-style-type: none"> • Zuordnung von DIF-Strings zu verlinkten GND-Orten über vorhandene GND-DIF-Paare • Unscharfer Vergleich von DIF- und GND-Strings
Berufe	<ul style="list-style-type: none"> • Konkordanzerstellung mit semantischer Übereinstimmung
Geschlecht	<ul style="list-style-type: none"> • Wertanpassung

Vorgaben für automatische Zusammenführung

Fall	Kurzbeschreibung	Fall	Kurzbeschreibung	Fall	Kurzbeschreibung
1a	1 Name + exakte Lebensdaten + 1 Beruf	3c	1 Name + GND-Namensvorkommen > 3 + exakte Lebensdaten	5a	1 Name + Lebensdaten Jahresübereinstimmung + Geburts- oder Sterbeort
1b	1 Name + exaktes Geburtsdatum oder exaktes Sterbedatum + 1 Beruf	4a	1 Name + GND-Namensvorkommen <= 3 + Lebensdaten Jahresübereinstimmung + 1 Beruf	5b	1 Name + Geburtsjahr oder Sterbejahr Jahresübereinstimmung + Geburts- oder Sterbeort
2a	1 Name + exakte Lebensdaten + Geburts- oder Sterbeort	4b	1 Name + GND-Namensvorkommen > 3 + Lebensdaten Jahresübereinstimmung + 1 Beruf	6	1 Name + GND-Namensvorkommen <= 3 + Lebensdaten Jahresübereinstimmung
2b	1 Name + exaktes Geburtsdatum oder exaktes Sterbedatum + Geburts- oder Sterbeort	4c	1 Name + GND-Namensvorkommen <= 3 + Geburtsjahr oder Sterbejahr + 1 Beruf	7	1 Name + GND-Namensvorkommen <= 3 + Geburts- oder Sterbeort + 1 Beruf
3a	1 Name + GND-Namensvorkommen <= 3 + exakte Lebensdaten	4d	1 Name + GND-Namensvorkommen > 3 + Geburtsjahr oder Sterbejahr + 1 Beruf	8	1 Name + Namensvorkommen = 1 + 2 Berufe aus filmportal
3b	1 Name + GND-Namensvorkommen <= 3 + exaktes Geburtsdatum oder exaktes Sterbedatum				

IN2N – laufender Betrieb

Vor Neuanlegen im DIF-Redaktionssystem wird automatisch geprüft, ob ein passender Datensatz in der GND vorhanden ist;

wenn ja -> Datensatz übernehmen

Wenn nein -> neuen Datensatz anlegen, der automatisch an die GND geliefert wird; GND-Nummer wird zurück geliefert und in den DIF-Datensatz geschrieben

Übertragung der Daten im Format GND/RDF

Bearbeiter: ge	Person	3B831252D784FE1A1C891100AB40692 [016525] DIF:P18552 CG:CBP020823 GND:123233216 GNDx:123233216
Login		Bearbeiten Ansichten: EAC-XML filmportal DNB-JSON
Übersicht		Monica Bleibtreu
Suchen	Namen	Monica Bleibtreu VN Monika Bleibtreu SV
Status-DB	Geschlecht	weiblich
KIM-DB	Geboren	04.05.1944 in Wien, Österreich
MIDAS	Gestorben	13.05.2009 in Hamburg
Wiki	Tätigkeit	Darstellerin, Drehbuch (1969 2009) in DE AT
Bearbeiter	Anmerkung	Auszeichnung: Deutscher Hörbuchpreis 2007 (18.03.2007), Hessischer Film- und Kinopreis: Ehrenpreis des Ministerpräsidenten (17.10.2008). <i>hat Tante oder Onkel: Hedwig Bleibtreu</i> <i>hat Geschwister: Renate Bleibtreu</i>

Datensatz aus der GND

Auszeichn	Aus: DNB Gemeinsame Normdatei (GND); ID: 123233216
	Vorzugsname: Monica Bleibtreu
	Weitere Namen: Monika Bleibtreu
	Geschlecht: female
	Geboren: 1944-05-04 in Wien
Filme	Gestorben: 2009-05-13 in Hamburg
	Berufe: Schauspielerin; Drehbuchautorin
	Biografisches: Auszeichnung: Deutscher Hörbuchpreis 2007 (18.03.2007), Hessischer Film- und Kinopreis: Ehrenpreis des Ministerpräsidenten (17.10.2008).
	Familiäre Beziehung zu: Bleibtreu, Moritz; Bleibtreu, Renate; Bleibtreu, Maximiliane; Bleibtreu, Renato Attilio
	Andere Quellen: Wikipedia: Monica_Bleibtreu; dbpedia: Monica_Bleibtreu; VIAF: 85128166
	<i>Azire über der neue Kontinent (1977/1978) Darsteller (Julie)</i>
	<i>Mittags auf dem Roten Platz (1977/1978) Darsteller</i>
	<i>Der Geist der Mirabelle. Geschichten von Bollerup (1978) Darsteller (Rita)</i>
	<i>Tatort: Alles umsonst (1978/1979) Darsteller (Anni Klein)</i>
	<i>Lemminge. 1. Arkadien. - 2. Verletzungen (1978/1979) Darsteller (Eva Beranek)</i>
	<i>Fallstudien (1979) Darsteller (Ina)</i>
	<i>Auf Achse: Fliegender Start (1980) Darsteller (Sylvia Mittermann)</i>
	<i>Auf Achse: Ganoven unter sich (1980) Darsteller (Sylvia Mittermann)</i>
	<i>Auf Achse: Nur eine kleine Verwechslung (1980) Darsteller (Sylvia Mittermann)</i>
	<i>Auf Achse: Vollgas (1980) Darsteller (Sylvia Mittermann)</i>
	<i>Alpensaga: Der deutsche Frühling (1980) Darsteller (Frau Leischner)</i>

Schnittstelle für partielle Datensatzänderungen

SCUSI = Simple Catalog Update Service Interface
(Arbeitstitel)

REST-basierte Webservice-Schnittstelle zum Erzeugen
oder Ändern von Katalogeinträgen

Drei Aktionen sind feldweise möglich:

- add
- replace
- remove

SCUSI - Nutzen

- Schnittstelle soll die automatische Übernahme von GND-Sätzen ermöglichen, wenn nicht der vollständige Datensatz benötigt wird / genommen werden kann
- Mitarbeit in der GND möglich, auch wenn nur Teile von Datensätzen angelegt / genutzt werden
- Kein Austausch der gesamten Datensätze nötig

SCUSI – weitere Planungen

- Schnittstelle derzeit für die Bedingungen des DIF eingerichtet
- Evaluierung nach einem halben Jahr
- Planung danach: Ausweitung der Festlegungen zur Nutzung der Schnittstelle, basierend auf den Anforderungen weiterer (möglicher) Partner

http://in2n.de/medien/2014/08/20140930_IN2N-Workshop_Haffner_IdeeProdukteKooperationsmodell.pdf

Danke für Ihre Aufmerksamkeit

Fragen?