

Alexander Haffner

Die GND und ihr Weg in das Semantic Web

Warum Normdaten?

- normierte Sucheinstiege zu bibliografischen Einheiten
 - erhöhte Benutzerfreundlichkeit bei der Recherche
- alles an einer Stelle
 - zentrale Datenverwaltung
 - einmalige Charakterisierung von Normdatenentitäten (Vorzugsbenennung, Verweisungsformen etc.)
- kooperative Erstellung und Nutzung
 - Daten sind allen Bibliotheken zugänglich

Die Gemeinsame Normdatei (GND)

Ausgangssituation - Getrennte Normdateien

- **PND**: Personennamendatei
 Personen der Formal- und Sacherschließung
- **GKD**: Gemeinsame Körperschaftsdatei
 Körperschaften der Formalerschließung
 (einschl. Gebietskörperschaften und Kongresse)
- **SWD**: Schlagwortnormdatei
 Schlagwörter (alle sekundär behandelten Entitäten)
 (einschl. Körperschaften, Geografika, Veranstaltungen,
 Werke, Schriftdenkmäler)
- **DMA-EST**: Einheitssachtitel musikalischer Werke
 (einschl. Expressionen)

Datensätze in Zahlen



Ausgangssituation: Getrennte Normdateien

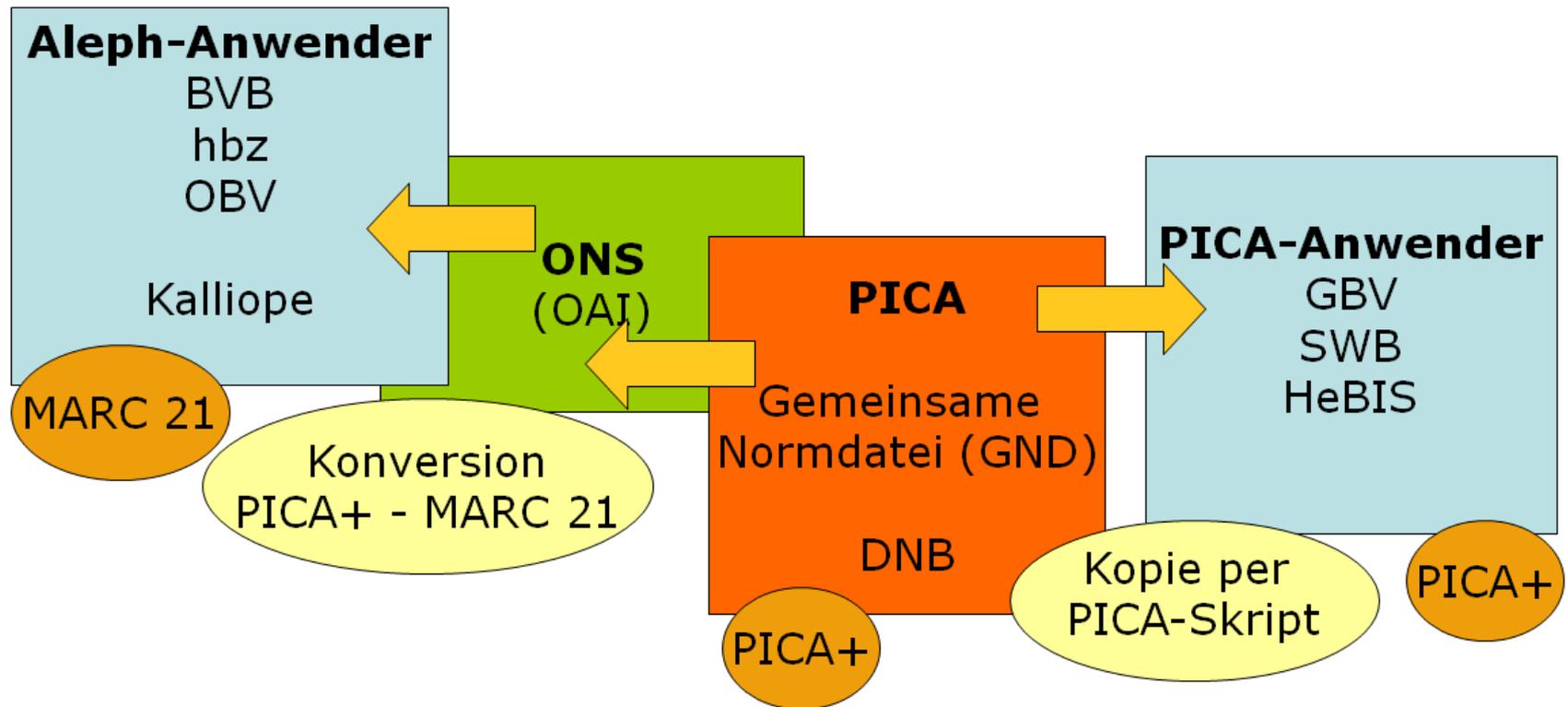
- unterschiedliche Regelwerke (RAK, RAK-Musik, RSWK)
- unterschiedliche Datenmodelle
- unterschiedliche Datenformate
- redundante Datenhaltung
- Anwender halten die Normdateien als Spiegel in der eigenen Arbeitsumgebung

Gemeinsame Normdatei (GND)

- Integration der Normdateien in objektorientierter Struktur
 - Entitätenbasierte Repräsentation
 - Aufhebung von Redundanzen
 - einheitliches Datenformat und Regelwerk
 - Bereinigung ungeeigneter Datenstrukturen
 - Möglichkeit der Eingabe von Originalschrift
- Angleichung an internationale Standards
 - Austauschformat MARC 21 – Authority
 - GND-Erfassungsformat nahe an MARC 21
 - Berücksichtigung von RDA, FRBR und FRAD
- Zukunftsfähig
 - Semantic Web-fähig
 - für automatische Verfahren geeignet

Kooperative Erschließung im Bibliothekswesen

Online-Normschnittstelle (ONS)



Online-Kommunikationsverfahren

- Eingabe von Datensätzen in die GND durch alle Teilnehmer im Bibliotheksverbund
- Erfassung und Änderung über den eigenen Client
 - „gleichzeitig“ in der überregionalen Normdatei und in der Verbunddatei
- alle Normdatensätze gelangen von Verbundebene in die überregionalen Normdateien
- neue Datensätze für alle Normdatenanwender über OAI verfügbar

Potentiale der semantischen Vernetzung

Ziele der semantischen Vernetzung

- Nachweis zusätzlicher Informationsquellen
- Vermeidung von Doppelarbeit bei der Erschließung
 - stärkere Nachnutzung externer Daten
 - Redundanzvermeidung, wenn möglich
- Aufmerksamkeit für bibliothekarische Normdaten steigern
- bessere Vernetzung der bibliografischen Daten
- Öffnung zu Nicht-Bibliothekscommunities

Semantische Vernetzung innerhalb des Bibliothekswesens

- Aufbau eines semantischen Netzes innerhalb der internationalen Bibliothekscommunity
- Voraussetzung:
 - Durchsuchbarkeit von Fremddaten möglichst ohne vorherigem Datenimport
- sofortige Einbeziehung in den Erschließungsprozess
 - aktuelle Praxis: nachgestellte Suche nach Konkordanzen/Redundanzen
- potentielle Datenquellen für die GND
 - VIAF, DDC, LCSH, STW etc.

Semantische Vernetzung mit externen Datenanbietern

- Engagement der Web 2.0 Community
- Überschneidungen bibliothekarischer Normdaten und Web 2.0 Plattformen identifizieren
- Ausweisung der Konkordanzen für den Katalogbenutzer
- Einbeziehung in die Erschließung (?)
 - Wo liegt die Kernkompetenz einer Bibliothek?
 - Gibt es Dinge, die andere besser machen?
 - Was ist heutzutage noch durch Bibliotheken zu leisten?

Konkordanzen zu Social Web Plattformen

- Wikipedia
 - Nachweis von Personen, Körperschaften, Geografika, Schlagwörtern, Werken
 - Konkordanz zu 162.492 Personen der PND
- GeoNames
 - Aggregator von Ämtern für Statistik und Vermessung
 - Nachweis von Geografika (sehr saubere Entitätentrennung)
 - Konkordanz zu 41.904 Orten und Ländern der SWD
- MusicBrainz
 - Nachweis von Personen und Körperschaften der Musik, Labels, Musiktonträgern
 - Konkordanz zu 31.322 Künstlern der PND und 14.663 der GKD

Stand November 2011

GND-Linked-Data-Service

Ziele

- Vereinfachung des Datenzugriffes für bibliotheksfremde Anwender
 - Zugriff über HTTP
 - kein vorheriger Datenimport notwendig
- Schaffung einer international nachnutzbaren GND-Datenrepräsentation im Web
 - dokumentierte Ontologie nutzbar durch Mensch und Maschine
- Vollständige Abbildung der Normdatenentitäten im Semantic Web
 - aktuell im Linked-Data-Service der DNB nur individualisierte Personen, Körperschaften und Sachschlagwörter
- verbesserte Vernetzung und Ausweisung von Konkordanzen

Der Weg zur GND im Semantic Web

- entitätenbasierte Repräsentation der GND-Daten
 - Personen, Familien, Körperschaften, Kongresse und Veranstaltungen, Geografika, Schlagwörter, Werke
- eine Ontologie zur Beschreibung der GND entwickeln
 - Voraussetzung:
 1. entitätenbasierter GND-Erfassungsleitfaden
 2. Beschreibung der verwendeten Felder und Codes
- Bereitstellung kontrollierter Vokabulare in RDF/SKOS
 - Sprachencodes, Ländercodes, DNB-Sachgruppen, Geschlechtsangaben, Funktionscodes etc.

Was machen andere?

<http://id.loc.gov/>

LIBRARY OF CONGRESS ASK A LIBRARIAN DIGITAL COLLECTIONS LIBRARY CATALOGS GO Options

The Library of Congress > Authorities & Vocabularies

AUTHORITIES & VOCABULARIES

Authorities and Vocabularies



- Authorities & Vocabularies
- About
- Vocabulary Descriptions
- Search
- Download
- Technical Center
- Contact Us

Search Authorities & Vocabularies

Hold CONTROL key for 2 or more

- All
- LC Subject Headings
- LC Name Authority File
- LC Children's Subject Headings
- LC Genre/Form Terms

Search

Available Authorities & Vocabularies

The Authorities and Vocabularies service provides access to commonly found standards and vocabularies promulgated by the Library of Congress. This includes data values and the controlled vocabularies that house them. The following are currently offered as part of this service:

- LC Subject Headings
- LC Name Authority File
- LC Children's Subject Headings
- LC Genre/Form Terms
- Thesaurus of Graphic Materials
- MARC Relators
- MARC Countries
- MARC Geographic Areas
- MARC Languages
- ISO639-1 Languages
- ISO639-2 Languages
- ISO639-5 Languages

Preservation Vocabularies

- Preservation File
- Preservation Level Role
- Cryptographic Hash Functions

Was machen andere?

<http://id.loc.gov/>

Available Authorities & Vocabularies

- [LC Subject Headings](#)
- [LC Name Authority File](#)
- [LC Children's Subject Headings](#)
- [LC Genre/Form Terms](#)
- [Thesaurus of Graphic Materials](#)
- [MARC Relators](#)
- [MARC Countries](#)
- [MARC Geographic Areas](#)
- [MARC Languages](#)
- [ISO639-1 Languages](#)
- [ISO639-2 Languages](#)
- [ISO639-5 Languages](#)

Preservation Vocabularies

- [Preservation Events](#)
- [Preservation Level Role](#)
- [Cryptographic Hash Functions](#)

MADS/RDF Namespace Document

- <http://www.loc.gov/standards/mads/rdf/v1.html>

Die GND-Ontologie

- Verwendung von OWL
- Definition der Entitätentypen als Klassen
- Differenzierung nach Datatype- und Object-Properties
 - Festlegung für Object-Properties ob asymmetric, transitive etc.
- Zuweisung von Domain und Range
 - vorerst nur Zuweisung der obersten Entitätenebene
- Definition von Value Vocabularies
- Zuordnung von Classes, Properties und Values anderer Ontologien (Vocabulary-Alignment)

Beispiel Datatype-Property

```

<owl:DatatypeProperty rdf:about="http://d-nb.info/gnd/schema/biographicalOrHistoricalInformation">
  <!-- Benennung -->
  <rdfs:label xml:lang="de">Biografische oder historische Angaben</rdfs:label>
  <rdfs:label xml:lang="en">Biographical or historical information</rdfs:label>
  <rdfs:comment xml:lang="de">Zusammenfassung von wesentlichen biografischen, geschichtlichen oder
    anderen Informationen über die zu beschreibende Entität.</rdfs:comment>
  <rdfs:comment xml:lang="en">Summary of the essential biographical, historical, or other
    information about the described entity.</rdfs:comment>
  <!-- MARC 21 Entsprechung -->
  <marc21equivalent>678 $a</marc21equivalent>
  <!-- Domain(s) & Range(s) -->
  <rdfs:domain rdf:resource="http://d-nb.info/gnd/schema/ConferenceOrEvent"/>
  <rdfs:domain rdf:resource="http://d-nb.info/gnd/schema/CorporateBody"/>
  <rdfs:domain rdf:resource="http://d-nb.info/gnd/schema/DiffereniatedPerson"/>
  <rdfs:domain rdf:resource="http://d-nb.info/gnd/schema/Family"/>
  <rdfs:domain rdf:resource="http://d-nb.info/gnd/schema/Geographica"/>
  <rdfs:domain rdf:resource="http://d-nb.info/gnd/schema/Subject"/>
  <rdfs:domain rdf:resource="http://d-nb.info/gnd/schema/Work"/>
  <rdfs:range rdf:resource="http://www.w3.org/2000/01/rdf-schema#Literal"/>
  <!-- Vocabulary Alignment -->
  <reg:superPropertyOf rdf:resource="http://rdvocab.info/ElementsGr2/biographicalInformation"/>
  <reg:superPropertyOf rdf:resource="http://rdvocab.info/ElementsGr2/familyHistory"/>
  <reg:superPropertyOf rdf:resource="http://rdvocab.info/ElementsGr2/corporateHistory"/>
  <rdfs:subPropertyOf rdf:resource="http://iflastandards.info/ns/fr/frad/P3039"/>
</owl:DatatypeProperty>
  
```

Beispiel Object-Property

```

<owl:ObjectProperty rdf:about="http://d-nb.info/gnd/schema/author">
  <!-- Benennung -->
  <rdfs:label xml:lang="de">Verfasserschaft</rdfs:label>
  <rdfs:label xml:lang="en">Author</rdfs:label>
  <!-- Beschreibung -->
  <rdfs:comment xml:lang="de">Eine Person, Familie oder Körperschaft, die verantwortlich für
    die Erstellung eines Werkes ist, das im Wesentlichen aus Text besteht, unabhängig vom
    Medientyp oder von der Gattung.</rdfs:comment>
  <rdfs:comment xml:lang="en">A person, family, or corporate body responsible for creating a
    work that is primarily textual in content, regardless of media type or genre.</rdfs:comment>
  <!-- MARC 21 Entsprechung -->
  <marc21equivalent>500 $4</marc21equivalent>
  <marc21equivalent>510 $4</marc21equivalent>
  <gndCode>auta</gndCode>
  <!-- Domain(s) & Range(s) -->
  <rdfs:domain rdf:resource="http://d-nb.info/gnd/schema/Work"/>
  <rdfs:range rdf:resource="http://d-nb.info/gnd/schema/CorporateBody"/>
  <rdfs:range rdf:resource="http://d-nb.info/gnd/schema/DifferentiatedPerson"/>
  <rdfs:range rdf:resource="http://d-nb.info/gnd/schema/Family"/>
  <!-- Vocabulary Alignment -->
  <rdfs:subPropertyOf rdf:resource="http://rdvocab.info/roles/author"/>
  <rdfs:subPropertyOf rdf:resource="http://purl.org/dc/terms/creator"/>
</owl:ObjectProperty>
  
```

Beispiel Object-Property

```

<owl:ObjectProperty rdf:about="http://d-nb.info/gnd/schema/gender">
  <!-- Benennung -->
  <rdfs:label xml:lang="de">Geschlecht</rdfs:label>
  <rdfs:label xml:lang="en">Gender</rdfs:label>
  <!-- Beschreibung -->
  <rdfs:comment xml:lang="de">Informationen zum Geschlecht einer Persn.</rdfs:comment>
  <rdfs:comment xml:lang="en">Information about gender of a person.</rdfs:comment>
  <!-- MARC 21 Entsprechung -->
  <marc21equivalent>375 $a</marc21equivalent>
  <!-- Domain(s) & Range(s) -->
  <rdfs:domain rdf:resource="http://d-nb.info/gnd/schema/DifferentiatedPerson"/>
  <rdfs:range rdf:resource="http://d-nb.info/gnd/vocabulary/genderValues"/>
  <!-- Vocabulary Alignment -->
  <owl:equivalentProperty rdf:resource="http://rdvocab.info/ElementsGr2/gender"/>
  <owl:equivalentProperty rdf:resource="http://xmlns.com/foaf/0.1/gender"/>
  <owl:equivalentProperty rdf:resource="http://iflastandards.info/ns/fr/frad/P3029"/>
</owl:ObjectProperty>
  
```

GND Gender

```
<skos:ConceptScheme rdf:about="http://d-nb.info/gnd/vocabulary/genderValues">
  <dc:title xml:lang="de">GND Geschlechterliste</dc:title>
  <dc:title xml:lang="en">GND Gender List</dc:title>
</skos:ConceptScheme>
```

```
<skos:Concept rdf:about="http://d-nb.info/gnd/vocabulary/genderValues#female">
  <!-- Scheme-Zuordnung -->
  <skos:inScheme rdf:resource="http://d-nb.info/gnd/genderValues"/>
  <!-- Benennung -->
  <skos:prefLabel xml:lang="en">female</skos:prefLabel>
  <skos:prefLabel xml:lang="de">weiblich</skos:prefLabel>
  <!-- Value Alignment -->
  <skos:exactMatch rdf:resource="http://rdvocab.info/termList/gender/1001"/>
</skos:Concept>
```

GND Country Code

```
<skos:ConceptScheme rdf:about="http://d-nb.info/gnd/vocabulary/countryCodes">
  <dc:title xml:lang="de">GND Ländercodes</dc:title>
  <dc:title xml:lang="en">GND Country Codes</dc:title>
</skos:ConceptScheme>
```

```
<skos:Concept rdf:about="http://d-nb.info/gnd/vocabulary/countryCodes#XA-DE">
  <!-- Scheme-Zuordnung -->
  <skos:inScheme rdf:resource="http://d-nb.info/gnd/countryCodes"/>
  <!-- Benennung -->
  <skos:prefLabel xml:lang="de">Deutschland</skos:prefLabel>
  <skos:prefLabel xml:lang="en">Germany</skos:prefLabel>
  <!-- Value Alignment -->
  <skos:closeMatch rdf:resource="http://id.loc.gov/vocabulary/countries/gw"/>
  <skos:closeMatch rdf:resource="http://d-nb.info/gnd/4011882-4"/>
</skos:Concept>
```

Weitere GND Value Vocabularies

Code Listen

- Language Code
- Content Type
- Carrier Type
- Media Type
- Form Of Work
- Function
- Instrument
- ...

Einführen einer Metametadata-Entität?

- Quelle und Datum der Ersterfassung
- Quelle und Datum der letzten Änderung
- Quelle und Datum der letzten Statusvergabe
- Katalogisierungslevel
- Kennzeichen für Hinweissatz
- Änderungs-codierung
- Teilbestandskennzeichen
- Nutzungskennzeichen
- Quellenangaben
- Negativ eingesehene Quellen
- Definition
- Benutzungshinweise
- Nummer und Ansetzungsform des Zielsatzes bei Umlenkung
- Katalogisierende Institution

Vielen Dank!

discussion is welcome...