

27. November 2018

Datendienst: Änderungen im Format RDF ab 12. Februar 2019 (Export-Release 2019.01)

1. Zeitliche Planung des Releases

Sehr geehrte Damen und Herren,

gemäß der Releaseplanung für Exportformate¹ der Deutschen Nationalbibliothek (DNB) wird die nächste Änderung in der Konversion und der Datenmodellierung am 12. Februar 2019 erfolgen. (Während der technischen Umstellung kann es an diesem Tag zwischen 8:00 – 12:00 Uhr zu Störungen der Schnittstellen und im Datashop kommen.)

2. Linked Data Service allgemein

2.1. Gesamtabzüge

In Anlehnung an die Exportreleases werden voraussichtlich in den Monaten Februar, Juni und November aktualisierte Gesamtabzüge der Gemeinsamen Normdatei (GND²) und der Titeldaten der DNB in den Serialisierungsformen RDF/XML, Turtle, JSON-LD, als HDT-Datei³ und N-Triples zum Download bereitgestellt⁴. Die Gesamtabzüge der Titeldaten der Zeitschriftendatenbank (ZDB) und der Adressdaten werden im März und im Oktober/November 2019 aktualisiert.

2.1.1. Provenienz-Informationen zu Titeldaten

Im Rahmen dieses Releases ist das Thema „Provenienz-Informationen zu den Titeldaten“ ein wichtiger Punkt, also Informationen zur Herkunft der einzelnen Tripel. Von Interesse ist dies vor allem für die Inhaltserschließungs-Elemente (dcterms:subject, dc:subject) in den Titeldaten. Es geht zum Beispiel um die Information, durch welchen Prozess ein GND-Schlagwort einer Publikation

¹ <https://wiki.dnb.de/x/wgcbBQ>

² <http://www.dnb.de/gnd>

³ Header, Dictionary, Triples; <http://www.rdfhdt.org/what-is-hdt/>

⁴ <https://data.dnb.de/opendata/>

zugeordnet wurde (intellektuell, maschinell mit Konfidenzwert XY, ...). Aber auch das formale Element „Sprachangabe“ (dcterms:language) ist ein entsprechender Anwendungsfall.

Diese Provenienz-Informationen werden entlang der PROV Ontology⁵ modelliert und zunächst voraussichtlich Ende November 2018 in Form eines separaten Datenabzugs veröffentlicht. Wir werden über die Mailingliste lds@lists.dnb.de entsprechend informieren und eine genaue Dokumentation der Modellierung bereitstellen.

2.2. In dem Dokument verwendete Präfixe

umbel:	http://umbel.org/umbel#
dcterms:	http://purl.org/dc/terms/
dc:	http://purl.org/dc/elements/1.1/
gndo:	http://d-nb.info/standards/elementset/gnd#
owl:	http://www.w3.org/2002/07/owl#
rdfs:	http://www.w3.org/2000/01/rdf-schema#

3. Veröffentlichung der RDF-Vokabulare

Angepasst an die Releaseplanung werden zusätzlich zu der Ankündigung der Änderungen im Format RDF auch datierte Vorab-Versionen der RDF-Vokabulare zu „DNB Metadata Terms“ und „GND Ontology“ erscheinen.

Für das Release 2019.02 sind die Vorab-Versionen abrufbar unter:

- | http://d-nb.info/standards/elementset/dnb_20190212 (DNB Metadata Terms)
- | http://d-nb.info/standards/elementset/gnd_20190212 (GND Ontology)

3.1. Änderungen in der GND-Ontologie

- | Korrektur: Die Hauptklassen der Gemeinsamen Normdatei waren bisher als sich gegenseitig ausschließend deklariert („owl:AllDisjointClasses“). Das betrifft „gndo:ConferenceOrEvent“, „gndo:CorporateBody“, „gndo:Family“, „gndo:PlaceOrGeographicName“, „gndo:SubjectHeading“ und „gndo:Work“. Diese Deklaration erwies sich als nicht korrekt, da gemäß den GND-Erfassungskonventionen Entitäten durchaus Instanzen mehrerer Klassen sein können.
Ein Beispiel: <http://d-nb.info/gnd/1099-6> („European Conference of Local Authorities“) ist in der Gemeinsamen Normdatei sowohl eine Kongressfolge („gndo:SeriesOfConferenceOrEvent“ ist „rdfs:subclassOf“ von „gndo:ConferenceOrEvent“) als auch eine Körperschaft („gndo:CorporateBody“). Aus diesem Grund wird das Axiom zunächst ersatzlos entfernt.

⁵ <https://www.w3.org/TR/prov-o/>

4. Änderungen in den GND-Konversionen

4.1. Profil GND-Ontologie

- | Korrektur: In sehr seltenen Fällen kam es bei der Bildung der Literale varianter Namensformen („gndo:variantNameForThePerson“) zu fälschlicher Stückelung mehrteiliger Namensangaben in mehrere Tripel. Dieser Fehler wurde jetzt behoben. Beispiel in der Testdatei: <http://d-nb.info/gnd/118806416>

Die folgenden Maßnahmen haben das Ziel, den Umfang der ausgelieferten Verlinkungen im RDF-Profil „GND-Ontologie“ an "Entity Facts" anzugleichen.

- | Neuerung: Aus dem Virtual International Authority File (VIAF⁶) werden zusätzlich weitere „owl:sameAs-Relationen“ übernommen. Diese sind im einzelnen:
 - o International Standard Name Identifier (ISNI⁷) („[http://www.isni.org/...](http://www.isni.org/)“)
 - o Wikidata⁸ („[http://www.wikidata.org/entity/...](http://www.wikidata.org/entity/)“)
 - o Library of Congress Authorities⁹ („[http://id.loc.gov/authorities/...](http://id.loc.gov/authorities/)“)
- | Verbesserung: Für den Entitätentyp „Person“ wird seit längerer Zeit, sofern vorhanden, eine Verlinkung mit „owl:sameAs“ auf die entsprechende VIAF-URI („[http://viaf.org/viaf/...](http://viaf.org/viaf/)“) ausgeliefert. Dies wird jetzt auf alle Entitätentypen der GND ausgeweitet.

4.2. Profil Entity Facts

- | Verbesserung: Bei der Modellierung der Geo-Koordinaten zu Entitäten des Typs "place" wurde zunächst kein RDF-Context angebunden¹⁰. Um die Entity-Facts-Daten durchgängig als JSON-LD abbilden zu können, werden nun Auszüge des GeoJSON-LD¹¹ als Context eingebunden. Da aktuell nur Punkt-Koordinatenangaben transportiert werden, ist das in dieser Form möglich. (In der GND ist eine Erfassung von Polygon-Koordinaten - die in GeoJSON nicht abbildbar sind - zwar theoretisch vorgesehen, aber praktisch nicht vorhanden.)
Beispiel in der Testdatei: alle
- | Neuerung: für Entitäten des Typs „place“ werden mit diesem Release auch „sameAs“-Verlinkungen auf <http://sws.geonames.org/> verfügbar.
Beispiel in der Testdatei: alle

⁶ <https://viaf.org/>

⁷ <http://www.isni.org/>

⁸ <https://www.wikidata.org>

⁹ <https://authorities.loc.gov/>

¹⁰ vgl. <https://lists.dnb.de/pipermail/lds/2018-May/000118.html>

¹¹ <http://geojson.org/geojson-ld/>

<http://hub.culturegraph.org/entityfacts/4018122-4>

```
{
  "@context" :
  "http://hub.culturegraph.org/entityfacts/context/v1/entityfacts.jsonld",
  "@type" : "place",
  "@id" : "http://d-nb.info/gnd/4018122-4",
  "preferredName" : "Frankfurt (Oder)",
  "sameAs" : [ {
    "@id" : "http://sws.geonames.org/2925535",
    "collection" : {
      "abbr" : "geonames",
      "name" : "GeoNames",
      "publisher" : "GeoNames",
      "icon" : "http://www.geonames.org/geonames.ico"
    }
  } ]
}
```

- | Verbesserung: Entity Facts¹² ist zukünftig auch verschlüsselt über das Protokoll HTTPS abrufbar. Es wird keine automatische Weiterleitung zwischen HTTP- und HTTPS-URLs eingerichtet - die Nutzer können sich aktiv für die HTTPS-Verbindung entscheiden. Wir empfehlen den Aufruf über HTTPS, da so die Authentizität der Entity-Facts-Daten und der Datenschutz für die Entity-Facts-einbindende Anwendung gewährleistet ist.

Die Contextdatei wird abhängig vom aufgerufenen URL über HTTP- oder HTTPS eingebunden. Die Beschreibung des Entity Fact-Dokuments bezieht sich fortan ausschließlich auf den HTTPS-URL als dem empfohlenen URL.

Ein Beispiel:

Der Aufruf von <https://hub.culturegraph.org/entityfacts/121966151> liefert folgende Daten (Auszug):

```
{
  "@context" :
  "https://hub.culturegraph.org/entityfacts/context/v1/entityfacts.jsonld",
  "@type" : "person",
  "@id" : "http://d-nb.info/gnd/121966151",
  "describedBy" : {
    "valid" : "2018-10-15T19:32:59+0200",
    "license" : "https://creativecommons.org/publicdomain/zero/1.0/",
    "@id" : https://hub.culturegraph.org/entityfacts/121966151
  }
}
```

¹² <http://www.dnb.de/entityfacts>

5. Änderungen in der Konversion der Titeldaten der Deutschen Nationalbibliothek und der Zeitschriftendatenbank

- | Änderung: Die Ausgabe des Persistent Identifiers vom Typ URN beginnt nun mit „[https://nbn-resolving.org/...](https://nbn-resolving.org/)“ anstatt „[http://nbn-resolving.de/...](http://nbn-resolving.de/)“. Es werden jedoch beide Resolveradressen zukünftig aktiv bleiben.

Beispiel in der Testdatei:

```
<http://d-nb.info/1162896434> umbel:isLike <https://nbn-resolving.org/urn:nbn:de:101:1-2018071705435047337751> .
```

- | Änderung: Die Ausgabe des Persistent Identifiers vom Typ DOI wird ebenfalls angepasst. Statt „[http://dx.doi.org/...](http://dx.doi.org/)“ beginnt der URL mit „[https://doi.org/...](https://doi.org/)“ .

Beispiel in der Testdatei:

```
<http://d-nb.info/1162896434> umbel:isLike <https://doi.org/10.1007/978-3-319-74983-9>
```

6. Beispieldateien und weitere Hinweise

Folgende Beispieldateien inklusive der angekündigten Änderungen sind in den Serialisierungsformen RDF/XML, Turtle und JSON abrufbar unter <https://data.dnb.de/testdat/>:

- | „GNDTestRelease2019_01 ...“ für die Normdaten der GND
- | „DNBTitelTestRelease2019_01 ...“ für die DNB-Titeldaten
- | „ZDBTitelTestRelease2019_01 ...“ für die ZDB-Titeldaten
- | „EntityFactsTestRelease2019_01.jsonld...“ (nur JSON-LD)

Informationen rund um den Linked-Data-Service der Deutschen Nationalbibliothek erhalten Sie auf unserer Webseite¹³.

Für Rückfragen stehen wir Ihnen gerne zur Verfügung.

Für Rückfragen stehen wir Ihnen gerne zur Verfügung.

Heike Eichenauer (Digitale Dienste)
und Jana Hentschke (Informationsinfrastruktur)

Telefon: +49-69-1525-1074

Telefax: +49-69-1525-1636

E-Mail: h.eichenauer@dnb.de

<http://www.dnb.de>

¹³ <https://www.dnb.de/lds>