

11. April 2019

Datendienste: Änderungen im Format RDF ab 18. Juni 2019 (Export-Release 2019.02)

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 18. Juni 2019 erfolgen. (Während der technischen Umstellung kann es an diesem Tag zwischen 8:00 – 12:00 Uhr zu Störungen bei den Schnittstellen und im Datenshop kommen.)

2. Linked Data Service allgemein

2.1. Gesamtabzüge

In Anlehnung an die Exportreleases werden in den Monaten März, Juni/Juli und voraussichtlich Oktober/November aktualisierte Gesamtabzüge der Gemeinsamen Normdatei (GND²) und der Titeldaten der DNB in den Serialisierungsformen RDF/XML, Turtle, JSON-LD und als HDT-Datei³ sowie als N-Triples zum Download bereitgestellt⁴. Der Gesamtabzug für Entity Facts⁵ erfolgt ebenfalls zu diesen Zeiten, jedoch nur im Format JSON-LD.

Die Gesamtabzüge der Titeldaten der Zeitschriftendatenbank (ZDB) und der Adressdaten werden jeweils im März und im Oktober/November aktualisiert.

¹ <https://wiki.dnb.de/x/wgcbBQ>

² <https://www.dnb.de/gnd>

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

⁴ <https://data.dnb.de/opendata/>

⁵ <https://www.dnb.de/entityfacts>

2.2. Provenienz-Informationen zu Titeldaten

Im Rahmen des letzten Releases war das Thema „Provenienz-Informationen zu den Titeldaten“ (Informationen zur Herkunft einzelner Tripel) ein wichtiger Punkt. 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 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 Ontology⁶ modelliert und in Form eines separaten Datenabzugs veröffentlicht. Dieser ist ab diesem Release auch in der Serialisierung JSON-LD verfügbar. Mit dieser Entwicklung konnte bereits ein Gesamtabzug Datenstand Release 2019.01 (13. Februar 2019) erstellt werden. Mehr Informationen zur Planung sowie die Dokumentation der Modellierung sind im WIKI⁷ der DNB verfügbar. Die Aktualisierung erfolgt jeweils in dem gleichen Rhythmus wie die Gesamtabzüge.

2.3. In dem Dokument verwendete Präfixe

dcterms:	http://purl.org/dc/terms/
gndo:	https://d-nb.info/standards/elementset/gnd#
owl:	http://www.w3.org/2002/07/owl#

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:

https://d-nb.info/standards/elementset/dnb_20190618 (DNB Metadata Terms)
https://d-nb.info/standards/elementset/gnd_20190618 (GND Ontology)

3.1. Änderungen:

- | Verbesserung: Alle RDF-Vokabulare wurden erweitert um die Aussagen „dcterms:license“⁸. Bisher war die Lizenz jeweils nur mit der Property „http://creativecommons.org/ns#license“ ausgewiesen.
- | Aufgrund von Nutzermeldungen weitere kleine Korrekturen

4. Änderungen in den GND-Konversionen

⁶ <http://www.w3.org/TR/prov-o/>

⁷ <https://wiki.dnb.de/x/46G5C>

⁸ <https://creativecommons.org/publicdomain/zero/1.0/>

- | Korrektur: Aus VIAF⁹ übernommene ISNI-URIs (über „sameAs“ im Profil Entity Facts und mit „owl:sameAs“ in Profil GND-Ontologie¹⁰) werden vom Präfix „http://www.isni.org/“ auf „http://isni.org/isni/“ korrigiert.¹¹
Beispiele in der Testdatei:
`http://d-nb.info/gnd/10000054-X`, `http://d-nb.info/gnd/100003141`,
`http://d-nb.info/gnd/10000324-2`
- | Korrektur: Aus VIAF übernommene -URIs (über „sameAs“ im Profil Entity Facts und mit „owl:sameAs“ in Profil GND-Ontologie¹²) werden vom Präfix „http://id.loc.gov/authorities/“ auf „http://id.loc.gov/rwo/agents/“ korrigiert.
Beispiel in der Testdatei:
`http://d-nb.info/gnd/118540238`

4.1. Profil GND-Ontologie

- | Korrektur: die Produktion von Aussagen mit der Property „gndo:playedInstrument“ (betrifft Entitäten des Typs Person) war versehentlich seit längerem inaktiv. Sie wurde jetzt wieder aktiv gesetzt.
Beispiel in Turtle:
`<http://d-nb.info/gnd/128910666> gndo:playedInstrument`
`<http://d-nb.info/gnd/7694384-7> .`
- Beispiele in der Testdatei:
`http://d-nb.info/gnd/134689240`, `http://d-nb.info/gnd/124883869`,
`http://d-nb.info/gnd/134739523`, `http://d-nb.info/gnd/128910666`,
`http://d-nb.info/gnd/128368497`

4.2. Profil Entity Facts

Neuerung: als weiteren Entitätentyp unterstützt Entity Facts jetzt „event“. Dabei handelt es sich um GND-Entitäten des Datensatztyps „Tf“, d.h. Konferenzen, Veranstaltungen, Konferenzfolgen bzw. -reihen, Veranstaltungsfolgen bzw. -reihen. Die einzelnen Beschreibungselemente und können der Gesamtübersicht JSON-LD¹³ entnommen werden.

Beispiel in JSON-LD:

```
{
  "@context" : "http://hub.culturegraph.org/entityfacts/context/v1/entityfacts.jsonld",
  "@type" : "event",
  "@id" : "http://d-nb.info/gnd/1161350764",
  "preferredName" : "Deutscher Bibliothekartag (107. : 2018 : Berlin)",
  "describedBy" : {
    "valid" : "2019-03-11T14:38:49+0100",
    "license" : "https://creativecommons.org/publicdomain/zero/1.0/",
    "@id" : "https://hub.culturegraph.org/entityfacts/1161350764"
  },
  "variantName" : [ "Bibliothekartag (107. : 2018 : Berlin)", "107. Deutscher Bibliothekartag" ],
  "dateOfEvent" : [ "12.06.2018-15.06.2018", "2018" ],
```

⁹ <https://viaf.org/>

¹⁰ vgl. WIKI Linked-Data-Service der DNB <https://wiki.dnb.de/x/TZa5C>

¹¹ vgl. <http://www.isni.org/how-isni-works>, Abschnitt „Linked data“

¹² vgl. WIKI Linked-Data-Service der DNB <https://wiki.dnb.de/x/TZa5C>

¹³ <https://wiki.dnb.de/x/9ob9Bw>

```

"organizerOrHost" : [ {
  "@id" : "http://d-nb.info/gnd/10014451-2",
  "preferredName" : "Berufsverband Information Bibliothek"
}, {
  "@id" : "http://d-nb.info/gnd/108169226X",
  "preferredName" : "VDB -Verein Deutscher Bibliothekarinnen und Bibliothekare"
} ],
"associatedPlace" : [ {
  "@id" : "http://d-nb.info/gnd/4011882-4",
  "preferredName" : "Deutschland"
} ],
"placeOfEvent" : [ {
  "@id" : "http://d-nb.info/gnd/4005728-8",
  "preferredName" : "Berlin"
} ],
"isA" : [ {
  "@id" : "http://d-nb.info/gnd/4130470-6",
  "preferredName" : "Kongress"
} ],
"topic" : [ {
  "@id" : "http://d-nb.info/gnd/4145267-7",
  "preferredName" : "Bibliothekerverband"
} ],
"isPartOf" : [ {
  "@id" : "http://d-nb.info/gnd/1024863-8",
  "preferredName" : "Deutscher Bibliothekartag"
} ],
"biographicalOrHistoricalInformation" : "Der 107. Deutsche Bibliothekartag fand unter dem Motto
\"offen & vernetzt\" im Tagungshotel Estrel statt",
"homepage" : [ "https://bibliothekartag2018.de/" ],
"sameAs" : [ {
  "@id" : "http://d-nb.info/gnd/1161350764/about",
  "collection" : {
    "abbr" : "DNB",
    "name" : "Gemeinsame Normdatei (GND) im Katalog der Deutschen Nationalbibliothek",
    "publisher" : "Deutsche Nationalbibliothek",
    "icon" :
"http://www.dnb.de/SiteGlobals/StyleBundles/Bilder/favicon.png?__blob=normal&v=1"
  }
}, {
  "@id" : "http://viaf.org/viaf/4353152988207812790008",
  "collection" : {
    "abbr" : "VIAF",
    "name" : "Virtual International Authority File (VIAF)",
    "publisher" : "OCLC",
    "icon" : "http://viaf.org/viaf/images/viaf.ico"
  }
} ]
}

```

Beispiele in der Testdatei:

<http://hub.culturegraph.org/entityfacts/1161350764>,

<http://hub.culturegraph.org/entityfacts/1010859-2>,

<http://hub.culturegraph.org/entityfacts/4032376-6>

<http://hub.culturegraph.org/entityfacts/1029455600>
<http://hub.culturegraph.org/entityfacts/5105778-5>
<http://hub.culturegraph.org/entityfacts/16211785-1>

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

- | Änderung: für Aussagen mit der Property „dcterms:accrualPeriodicity“ wird das verwendete RDF-Werte-Vokabular umgestellt. Betroffen sind vor allem ZDB-Ressourcen. Das bisher verwendete RDF-Werte-Vokabular „<http://marc21rdf.info/terms/continuingfre>“ wird abgelöst von „<http://id.loc.gov/vocabulary/frequencies>“, da dieses langfristig persistenter und stabiler eingeschätzt wird.

Beispiele in Turtle:

bisher:

```
<http://ld.zdb-services.de/resource/2018478-5> dcterms:accrualPeriodicity  
<http://marc21rdf.info/terms/continuingfre%23a> .
```

neu:

```
<http://ld.zdb-services.de/resource/2018478-5> dcterms:accrualPeriodicity  
<http://id.loc.gov/vocabulary/frequencies/ann> .
```

Beispiele in der Testdatei:

```
http://ld.zdb-services.de/resource/2018506-6,  
http://ld.zdb-services.de/resource/2018545-5  
http://ld.zdb-services.de/resource/2019636-2  
http://ld.zdb-services.de/resource/2018824-9
```

6. Vorausschau: Protokoll HTTPS für URIs und Permalinks

Voraussichtlich ab Mitte Oktober 2019 werden die Entity-URIs der Ressourcen der DNB, GND, ZDB und des ISIL- und Sigelverzeichnisses geändert. Das Protokoll HTTP wird ersetzt durch HTTPS. Das bedeutet, dass sich die URIs in den Subjektpositionen der RDF-Tripel ändern.

Beispiele:

<http://d-nb.info/gnd/134689240> -> <https://d-nb.info/gnd/134689240>

<http://d-nb.info/454884451> -> <https://d-nb.info/454884451>

<http://ld.zdb-services.de/resource/2018506-6> -> <https://ld.zdb-services.de/resource/2018506-6>

<http://ld.zdb-services.de/resource/organisations/DE-26-010> -> <https://ld.zdb-services.de/resource/organisations/DE-26-010>

Wir haben uns zu diesem Schritt entschlossen, um die Daten in allen Fällen über gesicherte Verbindungen auszuliefern und damit die Authentizität der Daten zu jedem Zeitpunkt gewährleisten

zu können. Die HTTP-URIs werden dauerhaft auf die neuen HTTPS-URIs weitergeleitet (HTTP-Statuscode „301 Moved Permanently“).

7. Beispieldateien und weitere Hinweise

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

- | „GNDTestRelease2019_02 ...“ für die Normdaten der GND
- | „ZDBTitelTestRelease2019_02 ...“ für die ZDB-Titeldaten
- | „EntityFacts_TestRelease2019_02.jsonld“ für Entity Facts

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

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

<mailto:h.eichenauer@dnb.de>

<https://www.dnb.de>