1. November 2018

Data service and interfaces: Machine assignment of keywords from the Library of Congress Subject Headings beginning for English-language university dissertations published online

Since June 2018, the German National Library (DNB) has been assigning keywords\(^1\) to all English-language university dissertations for which full-text versions are available (online publications\(^2\)) by machine. This relates to series O metadata of the Deutsche Nationalbibliografie\(^3\). In future, this information will also be enriched in the metadata for parallel issues in series A, B and H of the Deutsche Nationalbibliografie.

Topical terms are now assigned using Library of Congress Subject Headings (LCSH\(^4\)). For the proper names the vocabulary from the Integrated Authority File\(^5\) (GND entities: persons, geographical information, corporate bodies, conferences and works) will be used.

Delivery using export formats

1. **MARC 21 format**

Export release 2019.01\(^6\) on 12 February 2019 will contain the machine-assigned LCSH keywords in export format MARC 21\(^7\). Delivery will be effected in the known MARC fields 6XX (Subject Access Fields). The information “maschinell gebildet” (machine generated) will be entered in the corresponding MARC field 883 (Machine-Generated Metadata Provenance), subfield $a.

---

\(^1\) [http://www.dnb.de/subjectcataloguing](http://www.dnb.de/subjectcataloguing)

\(^2\) [http://www.dnb.de/EN/netzpublikationen](http://www.dnb.de/EN/netzpublikationen)

\(^3\) [http://www.dnb.de/EN/nationalbibliografie](http://www.dnb.de/EN/nationalbibliografie)

\(^4\) [http://id.loc.gov/authorities/subjects.html](http://id.loc.gov/authorities/subjects.html)

\(^5\) [http://www.dnb.de/EN/gnd](http://www.dnb.de/EN/gnd)

\(^6\) [https://wiki.dnb.de/x/IoiRAg](https://wiki.dnb.de/x/IoiRAg)

\(^7\) [http://www.dnb.de/EN/marc21](http://www.dnb.de/EN/marc21)
2. RDF format

The delivery of LCSH and the corresponding provenance information is planned for the same time as export release 2019.01 as part of the DNB’s Linked Data Service. The provenance information will be provided as a separate data copy in RDF for the data set “DNBTitel”. Further details will be provided in an announcement scheduled for November 2018.

Example in RDF format (Turtle serialisation):

```turtle
@prefix dnbt: <http://d-nb.info/standards/elementset/dnb#>.
@prefix prov: <http://www.w3.org/ns/prov#>.
@prefix xsd: <http://www.w3.org/2001/XMLSchema#>.
@prefix agrelon: <http://d-nb.info/standards/elementset/agrelon#>.
@prefix dnb-prov-a: <http://d-nb.info/provenance/activity/>.

<http://d-nb.info/115505668X> dnbt:hasqualifiedSubject [ a prov:Entity, dnbt:QualifiedSubject ;
dnbt:qualifiesSubject <http://id.loc.gov/authorities/subjects/sh85134146> ;
prov:generatedAtTime "2018-05-22"^^xsd:date ;
agrelon:metadataConfidence "0.18692"^^xsd:float ;
prov:wasGeneratedBy dnb-prov-a:aep-lcsh:2018-05-22 ].
```

From export release 2019.02, the machine-generated LCSH keywords – excluding provenance information – will also appear in the regular profile “DINI-AG KIM recommendations”:

```turtle
@prefix dcterms: <http://purl.org/dc/terms/>.

```

Display in DNB catalogue

It is expected that the machine-assigned subject headings will be displayed in the full bibliographic information provided in the catalogue from 2019. Example: http://d-nb.info/1165650355

---

8 http://www.dnb.de/EN/lds
9 http://www.dnb.de/dataservicesnews