

12. Februar 2018

Datendienst Änderungen im Format MARC 21 für Titeldaten ab 15. Mai 2018 (Export-Release 2018.02)

Sehr geehrte Damen und Herren,

wir möchten Sie auf das Export-Release 2018.02 des Formats MARC 21 für Titeldaten aufmerksam machen.

1. Zeitplan

1.1. Externe Testphase

Im Rahmen der Releaseplanung¹ existiert eine externe Testphase. Zu Beginn dieser Phase erfolgt die Veröffentlichung der Feldbeschreibung und der Testdaten (s. Abschnitt 4). In den folgenden zwei Wochen können alle Datenbezieher die Testdaten im Hinblick auf Fehler oder Auffälligkeiten begutachten und diese melden.

Für das Release 2018.02 ist die externe Testphase:

| 12. – 23. Februar 2018

Rückmeldungen inkl. Beispiel bitte bis zum 23. Februar 2018 an datendienste@dnb.de.

Nach Ablauf der externen Testphase werden keine weiteren Anpassungen am Release vorgenommen.

Die Testdaten stehen bis zur Produktivnahme des Releases zum Testen des Imports zur Verfügung.

1.2. Produktivnahme

Die Änderungen sind gültig ab:

| 15. Mai 2018 (Während der technischen Umstellung kann es zwischen 8:00 – 12:00 Uhr zu Störungen bei den Schnittstellen und im Datenshop kommen.)

¹ <https://wiki.dnb.de/x/wgcbBQ>

2. Formatänderungen

2.1. MARC-Feld 034 (Kodierte kartografische mathematische Angaben) (R), neues Unterfeld \$a (Kategorie des Maßstabs) (NR)

Zur Vervollständigung und Präzisierung der Angaben im MARC-Feld 034 wird das Feld beim Export um das Unterfeld \$a erweitert. Als Standardangabe für die Kategorie des Maßstabs wird immer der Wert „a“ für „Linearmaßstab“ angegeben.

MARC-Feld 034 (Kodierte kartografische mathematische Angaben) (R)

- | Indikator 1 (Art des Maßstabs):
„0“ - Nicht definierter Maßstab/Kein Maßstab erfasst (= Default)
- | Indikator 2 (Art des Ringes):
„_“ - Nicht anwendbar (= Default)
„0“ - Äußerer Ring
„1“ - Ausschließender Ring
- | Neues Unterfeld:
\$a (Kategorie des Maßstabs) (NR) – „a“ (Linearmaßstab)

Beispiel in Auszügen:

034 0_ \$aa\$d W 118 14 37\$e W118.243611\$f N034.052222\$g N034.052222\$9A: dgx

2.2. MARC-Feld 041 (Sprachcode) (R), neues Unterfeld \$8 (Feldverknüpfung und Reihenfolge) (R) bei maschinell ermittelten Sprachen und entsprechende Ergänzung zu den Provenienzinformatoren im verknüpften MARC-Feld 883

Für Online-Ressourcen wird die maschinell ermittelte Sprache der Vorlage im MARC-Feld 041, Unterfeld \$a (Sprachcode des Textes / der Tonspur oder des separaten Titels) (R) mit dem entsprechenden Sprachcode² nach ISO 639-2 ausgeliefert. Mittelfristig ist geplant, mehrere maschinell ermittelte Sprachcodes auszuliefern und diese in jeweils eigenen Unterfeldern \$a auszuliefern.

Neu ist die Auslieferung des Unterfelds \$8 (Feldverknüpfung und Reihenfolge) (R) für die satzinterne Verknüpfung zum MARC-Feld 883 (Herkunft von maschinell generierten Metadaten) (R), das weitere Angaben zur Herkunft transportiert. Der Konfidenzwert für den Sprachencode enthält i.d.R. fünf Nachkommastellen.

² <http://www.loc.gov/marc/languages/>

MARC-Feld 041 (Sprachcode) (R)

- | Indikator 1 (Übersetzungshinweis):
„_“ – Keine Angabe vorhanden
- | Indikator 2 (Quelle des Codes):
„_“ – MARC-Sprachcode
- | Unterfelder:
\$a (Sprachcode des Textes/der Tonspur oder des separaten Titels) (R)-
MARC-Sprachcode, „ger“ = Default
\$c (Sprache des Originals) (R) – Sprachcode³
\$8 (Feldverknüpfung und Reihenfolge) (R) – Zählziffer, „\p“ (Verknüpfungstyp
„Metadata provenance“)

MARC-Feld 883 (Herkunft von maschinell generierten Metadaten) (R)

- | Indikator 1:
„0“ – vollständig maschinell generiert
- | Indikator 2:
„_“ – Nicht definiert
- | Unterfelder:
\$a Generierungsprozess – „maschinell gebildet“
\$c Konfidenzwert – „1,00000“ – „0,00000“
\$d Generierungsdatum – „JJJJMMTT“
\$q Generierungsagentur – „DE-101“
\$8 Feldverknüpfung und Reihenfolge – Zählziffer, „\p“ (z. B. „1\p“)
\$8 ist immer das erste Unterfeld.

Beispiel in Auszügen:

041 __ \$81\p\$aeng

883 0_ \$81\p\$amaschinell gebildet\$c4,91000\$d20171024\$qDE-101

2.3. MARC-Felder 082 (Notation nach der Dewey Decimal Classification) (NR) und 083 (Zusätzliche DDC-Notation) (R): Auslieferung maschinell erstellter DDC-Kurznotationen und entsprechende Ergänzung zu den Provenienzinformatoren im verknüpften MARC-Feld 883

Die Auslieferung maschinell gebildeter Sacherschließungselemente wird ergänzt um die Auslieferung von automatisiert vergebenen DDC-Notationen (Kurznotationen). Die erste und einzige Kurznotation wird entweder im MARC-Feld 082 (Notation nach der Dewey Decimal Classification) (NR) oder im MARC-Feld 083 (Zusätzliche DDC-Notation) (R) ausgeliefert, abhängig davon, ob eine weitere **vollständige** DDC-Notation vorliegt. Bei der Besetzung des nicht-wiederholbaren Feldes 082 wird folgende Reihenfolge berücksichtigt:
Vollständige DDC-Notation > DDC-Kurznotation > DDC-Sachgruppe.

³ Codes s. <http://www.loc.gov/marc/languages>

Die Herkunftsangaben der erstmalig ausgelieferten maschinell gebildeten Kurznotationen werden wie üblich im MARC-Feld 883 (Herkunft von maschinell generierten Metadaten) (R) verknüpft.

Die DDC-Kurznotation wurde zunächst nur für die Hauptsachgruppe 610⁴ (Medizin, Gesundheit) der Online-Ressourcen der „Reihe O“ der Deutschen Nationalbibliografie maschinell vergeben. Seit September 2017 werden die medizinischen Veröffentlichungen der Reihen B und H der Deutschen Nationalbibliografie ebenfalls mit diesen Kurznotationen erschlossen. Sachgruppe 004 (Informatik), 300 (Sozialwissenschaften) und 540 (Chemie) folgen demnächst. Für alle anderen Fächer wird zurzeit an einem Klassifikationsschema mit verkürzten DDC-Notationen gearbeitet. Geplant ist, dass die entsprechenden Kurznotationen in WebDewey Deutsch sowie im frei zugänglichen Retrievaltool WebDewey Search gekennzeichnet werden.

MARC-Feld 082 (Notation nach der Dewey Decimal Classification) (NR)

- | Indikator 1 (Art der Ausgabe):
„7“ – Andere Ausgabe näher beschrieben im Unterfeld \$2
- | Indikator 2 (Quelle der Notation):
„4“ – Von anderer Stelle als der LC vergeben
- | Unterfelder:
\$a (Notation) (R)- DDC-Kurznotation
\$q (Vergabestelle) (NR) – „DE-101“ Deutsche Nationalbibliothek
\$2 (Ausgabennummer) (NR) - „23kdnb“ (neu)
\$8 (Feldverknüpfung und Reihenfolge) (R) – Zählziffer „\p“ (Verknüpfungstyp „Metadata provenance“)

MARC-Feld 083 (Zusätzliche DDC-Notation) (R) wird wie MARC-Feld 082 besetzt; Ausnahme: Der 2. Indikator ist immer ein „_“ (Blank).

MARC-Feld 883 (Herkunft von maschinell generierten Metadaten) (R)

s. Abschnitt 2.2

Beispiel in Auszügen:

082 74 \$84\p\$a616.2\$qDE-101\$223kdnb

883 0_ \$84\p\$amaschinell gebildet\$c0,758\$d20170308\$qDE-101

3. Weitere Hinweise

- | MARC-Felder 76X – 78X: Die im Unterfeld \$z gelieferte ISBN wird durchgängig ohne Bindestriche geliefert.
- | MARC-Leader Position 07 (Bibliografischer Level): Die Besetzung dieser Position mit dem Wert „i“ (Integrierende Ressource) wurde für fortlaufende Ressourcen erweitert. Sind die Quelldaten mit den Inhaltstypen „Datenbank“, „Website“, „Loseblattsammlung“ oder „Weblog besetzt“, wird ein „i“ ausgeliefert.

⁴ <http://www.dnb.de/Subsites/ddcdeutsch/SharedDocs/Downloads/DE/anwendung/ddcGliederungMedizin.pdf>

- | MARC-Feld 008 CR (Feld mit fester Länge zur physischen Beschreibung - Fortlaufende Ressourcen), Position 21 (Art der fortlaufenden Ressource): Die Besetzung mit den Werten „d“ und „w“ wurde ebenfalls neu strukturiert. Der Wert „d“ (fortwährend aktualisierte Datenbank) wird gesetzt, wenn der Inhaltstyp „Datenbank“ bei Zeitschriften vorliegt und der Wert „w“ (fortwährend aktualisierte Webseite) wird gesetzt, wenn die Inhaltstypen „Website“ oder „Weblog“ verknüpft sind.
- | MARC-Feld 255 (Kartografische mathematische Daten) (R): mehrere Maßstabsangaben werden jeweils in einem eigenen MARC-Feld 255 ausgeliefert.
- | MARC-Feld 650 (Nebeneintragung unter einem Schlagwort, Sachschlagwort): Mehrere Schlagwörter aus Fremddaten ohne Verknüpfung zur GND, die in einem Unterfeld \$a ausgeliefert werden, werden nun in der Regel durch „;“ voneinander getrennt, wenn die Struktur der Quelldaten dies ermöglicht.

4. Weiterführende Informationen und Testdaten

Die offizielle Dokumentation der beschriebenen Felder finden Sie auf den Webseiten der Library of Congress unter <http://www.loc.gov/marc/bibliographic/ecbdhome.html>.

Die vollständige Feldbeschreibung der Titeldaten ist als PDF- und Excel-Version hinterlegt unter <http://www.dnb.de/marc21>.

Um einen schnellen Überblick über die Änderungen seit der Veröffentlichung der vorhergehenden Beschreibung anzubieten, sind in der PDF-Version alle Änderungen gegenüber der unmittelbar vorhergehenden Version auf den ersten Seiten vorangestellt.

Eine Beispieldatei, die die angekündigten Änderungen enthält, finden Sie unter <https://data.dnb.de/testdat/> mit dem Dateinamen „TitelTestRelease2018_02“.

5. Unverbindliche Vorabinformation für das übernächste Release

Bitte beachten Sie, dass diese Vorabinformation auf mögliche zukünftige Änderungen nicht vollständig sein kann, auch können Änderungen auf ein späteres Release verschoben werden oder ganz entfallen. Ebenso können sich noch wesentliche inhaltliche Änderungen ergeben. Welche Punkte konkret umgesetzt werden, entnehmen Sie bitte der entsprechenden Ankündigung des Releases, die Anfang Juli 2018 versendet wird.

Weitere Informationen zum Release-Zyklus unter: <https://wiki.dnb.de/x/wgcbBQ>

Für das **übernächste Release** Anfang Oktober 2018 sind folgenden Änderungen in der Planung:

Im Bereich Titeldaten:

- | Auslieferung von maschinell vergebenen englischsprachigen Schlagwörter⁵ der LCSH in den MARC-Feldern 6XX sowie der entsprechenden Herkunftsangaben im MARC-Feld 883.
- | Korrektur in den Verknüpfungsfeldern 76X – 78X: in bestimmten seltenen Konstellationen werden falsche MARC-Felder ausgeliefert. Dies wird korrigiert.
- | Weitere Anpassung in den Verknüpfungsfeldern 76X – 78X für den Export von Angaben zu verknüpften Werken, die nicht in der ZDB vorhanden sind.

⁵ <http://www.loc.gov/aba/publications/FreeLCSH/freelcsh.html>

- | Für Sammlungen wird im Leader an Position 07 (Bibliografischer Level) ein „c“ für Sammlung geliefert. Für diese Datensätze wurde bisher im MARC-Feld 008 an Position 06 (Art der Datumsangabe/Veröffentlichungsstatus) ein „c“ (nicht abgeschlossene fortlaufende Ressource) oder ein „d“ (abgeschlossene fortlaufende Ressource) geliefert. Hier sollen diese Werte nun durch ein „i“ (Entstehungszeitraum einer Sammlung) ersetzt werden.

Für weitere Fragen erreichen Sie uns per E-Mail unter datendienste@dnb.de und unter den bekannten Telefonnummern.

Mit freundlichen Grüßen

i. A. Heike Eichenauer

*** Lesen. Hören. Wissen. Deutsche Nationalbibliothek ***

--

Heike Eichenauer
Deutsche Nationalbibliothek
Digitale Dienste
Adickesallee 1
D-60322 Frankfurt
Telefon: +49-69-1525-1074
Telefax: +49-69-1525-1636
<mailto:h.eichenauer@dnb.de>
<http://www.dnb.de>