29 June 2020

Metadata services: changes in MARC 21 format for bibliographic data from 29 September 2020 (Export Release 2020.03)

Dear Sir or Madam

We would like to draw your attention to the Export Release 2020.03 of the MARC 21 format for bibliographic data.

1. Schedule

1.1. External test phase

An external test phase is included in the release planning[1]. The field description and test data will be made public at the beginning of this phase (see section 4). During the first two weeks following publication, all data subscribers will be able to examine the data for errors or anomalies and report them as necessary.

For the 2020.03 release, the external test phase is:

> | 29 June – 10 July 2020

Please send any feedback (including examples) by 10 July 2020 to metadatendienste@dnb.de.

No further adjustments will be made to the scheduled release once the external test phase has ended. If we receive feedback at a later date, we will examine whether any corrections can be made; such corrections will not be issued any earlier than the following release.
After the end of the external test phase, the test data will be available for testing the import until the release goes into productive use.

---

[1] https://wiki.dnb.de/x/wgcbBQ

1.2. Productive use

The changes will apply from:

| 29 September 2020 (There may be disruptions to the interfaces and the Data Shop between 08:00 and 12:00 noon during the technical changeover.)

2. Format changes

2.1. General revision of MARC field 883 (Metadata provenance)

Update 30[2] to the MARC documentation included the expansion of MARC field 883 (Metadata provenance), which is now also able to transport statements regarding the provenance of data generated intellectually. MARC field 883 was substantially revised for this reason. In the case of bibliographic data which contains information on the way in which it was generated (intellectually (new value "2" for the first indicator), by machine, or by data import, mapping etc.), information on the creation process is issued and a URI delivered which references the process description on our website. The information in subfield $a is now only transmitted in coded form. The contents of $c (Confidence value) and $d (Creation date) are delivered unchanged.

2.1.1. New value "2" (Not machine-generated) for the 1st indicator

1st indicator (Method of assignment):

| # No information provided/not applicable
| 0 Fully machine-generated
| 1 Partially machine-generated
| 2 Not machine-generated

2.1.2. New values in $a (Creation process) (NR)

The description of the creation process in subfield $a is no longer delivered in text form but in the form of a code. See also subfield $u for a more detailed process description.

Examples of a number of codes:

| **dnb** Generated intellectually by the DNB
| **cgwrk** Machine-based import from Culturegraph work bundle
| **aeplc** Machine-based assignment of language codes
| **npi** Taken over by the DNB when importing online publications
| **vlb** Imported by the DNB from MVB's German Books in Print

---

[2] https://www.loc.gov/marc/bibliographic/bdapndxg.html

| **orcid**   Saved by the person themselves in the ORCID record and imported by
                     the DNB

2.1.3. New subfield $u (Uniform Resource Identifier) (NR)

The new subfield $u delivers a URI in the form https://d-nb.info/provenance/plan#CODE, which references the process description on our website. The table[3] is completed and continually updated until the release will go live.

It is expected that legacy data containing provenance information will continue to exist until the end of 2020. This means that subfields $a and $u could refer to processes that are actually no longer valid and are therefore not described in the table. If this is the case, the URI in $u will not resolve, and unfortunately be unable to deliver the desired result. However, a clean-up of obsolete data is being planned.

MARC field 883

First indicator (Method of assignment):

| „_"   – No information provided/not applicable
| „0"   – Fully machine-generated
| „1"   – Partially machine-generated
| „2"   – Not machine-generated

Second Indicator (Undefined):

| „_"   – Undefined

Subfields:

| $a (Creation process) (NR):   Code describing the creation process
                                                 (see also $u)
| $c (Confidence value) (NR):   "1.00000" - "0.00000" (if available)
| $d (Creation date)   "YYYMMDD" – date of automatic creation
| $q (Assigning or creating agency) (NR)   "DE-101"
| $u (Uniform Resource Identifier) (NR)   URI of process description in the form
                                                 https://d-nb.info/provenance/plan#Code
| $8 (Field link and sequence number) (R)   Numerical code for internal link "\p"
                                                 (link type "Metadata provenance")
                                                 Example: "1\p"
                                                 $8 is always the first subfield.

2.2. Various MARC fields: New $8 subfields

Due to the revision of MARC field 883, other MARC fields containing a subfield $8 will have the code "\p" attached as the field link type.

---

[3] https://d-nb.info/provenance/plan

The following MARC fields have a \p link in subfield $8:

	041 (R)	Language Code
	082 (NR) / 083 (R)	DDC, DDC Short numbers, DDC Subject categories
	084 (R)	Other Notations
	100 (NR) / 700 (R)	Main entry and added entry personal name
	110 (NR) / 710 (R)	Main entry and added entry corporate name
	111 (NR) / 711 (R))	Main entry and added entry meeting name
	385 (R)	Audience characteristics
	600-655 (R) / 689 (R)	Subject added entries (RSWK)
	600-651 (R)	Subject added entries (machine-generated subject entries, external subject entries)
	653 (R)	Index term - uncontrolled
	655 (R)	Index term – genre/form
	883 (R)	Metadata provenance

Example (simulated, version following clean-up of obsolete data):

041 __ $8\1\p$ager$aeng

041 __ $8\2\p$aita

082 04 $8\12\p$a349.494542409$qDE-101$223/ger

083 7_ $8\14\p$a618.92$qDE-101$223kdnb

083 7_ $8\15\p$a610$qDE-101$222sdnb

084 __ $8\13\p$aFliegenpilz$2Waldthesaurus

100 1_ $8\5\p$0(DE-588)142347310$0https://d-nb.info/gnd/142347310$0(DE-101)142347310$aGraf, Tanja$eVerfasser$4aut$2gnd

385 __ $8\4\p$0(DE-588)4035088-5$0https://d-nb.info/gnd/4035088-5$0(DE-101)040350886$aLehrer$2gnd

650 _7 $8\7\p$0(DE-588)4038108-0$0https://d-nb.info/gnd/4038108-0$0(DE-101)040381080$aMaya$2gnd

650 _7 $8\8\p$0(DE-588)4062305-1$0https://d-nb.info/gnd/4062305-1$0(DE-101)04062305X$aValidität$2gnd

650 _7 $8\9\p$0(DE-588)4062862-0$0https://d-nb.info/gnd/4062862-0$0(DE-101)040628620$aVerhaltensforschung$2gnd

650 _7 $8\10\p$0(DE-588)4185497-4$0https://d-nb.info/gnd/4185497-4$0(DE-101)041854977$aTierpsychologie$2gnd

653 __ $8\11\p$a(Thesaurus)Deskriptor

655 _7 $8\3\p$0(DE-588)4048476-2$0https://d-nb.info/gnd/4048476-2$0(DE-101)040484769$aRatgeber$2gnd-content

710 2_ $8\6\p$0(DE-588)1065135297$0https://d-nb.info/gnd/1065135297$0(DE-101)1065135297$aVerlag Vorwerk 8$eVerlag$4pbl$2gnd8

883 1_ $8\1\p$avlb$d20200617$qDE-101$uhttps://d-nb.info/provenance/plan#vlb

883 2_ $8\2\p$adnb-pa$d20200617$qDE-101$uhttps://d-nb.info/provenance/plan#dnb-pa

883 __ $8\3\p$acgwrk$d20200617$qDE-101$uhttps://d-nb.info/provenance/plan#cgwrk

883 1_ $8\4\p$avlb$d20200617$qDE-101$uhttps://d-nb.info/provenance/plan#vlb

883 1_ $8\5\p$aidtitel$d20200617$qDE-101$uhttps://d-nb.info/provenance/plan#idtitel

883 1_ $8\6\p$aadb$d20200617$qDE-101$uhttps://d-nb.info/provenance/plan#adb

883 0_ $8\7\p$aaepgnd$c0,08373$d20200617$qDE-101$uhttps://d-nb.info/provenance/plan#aepgnd

883 0_ $8\8\p$aaepgnd-pa$c0,05617$d20200617$qDE-101$uhttps://d-nb.info/provenance/plan#aepgnd-pa

883 __ $8\9\p$agndddc$c1$d20200617$qDE-101$uhttps://d-nb.info/provenance/plan#gndddc

883 __ $8\10\p$agndddc$c1$d20200617$qDE-101$uhttps://d-nb.info/provenance/plan#gndddc

883 __   **$8**11\p**$a**cgwrk**$d**20200617**$q**DE-101**$u**https://d-nb.info/provenance/plan#cgwrk
883 __   **$8**12\p**$a**cgwrk**$d**20200617**$q**DE-101**$u**https://d-nb.info/provenance/plan#cgwrk
883 __   **$8**13\p**$a**cgwrk**$d**20200617**$q**DE-101**$u**https://d-nb.info/provenance/plan#cgwrk
883 0_   **$8**14\p**$a**aepkn**$c**0,99481**$d**20200617**$q**DE-101**$u**https://d-nb.info/provenance/plan#aepkn
883 0_   **$8**15\p**$a**aepsg**$c**0,97199**$D**2020-06-17**$q**DE-101**$u**https://d-nb.info/provenance/plan#aepsg

3. Other information

| MARC field 090 (Other codes) (NR) $a (Paper condition) (NR):
From 1 October 2020, the code "b" for "Documents printed on acid-free or low-acid paper" will no longer be delivered with current data. This information will still be provided in legacy data.

| MARC field 880 (Alternate graphic representation) (R), subfield $2 (Source) (NR):
The subfield $2 has now been added to MARC field 880 with code "gnd" . This is part of a general bug-fixing concerning MARC fields that link to the GND and the declaration "gnd" in the subfield $2. The information in MARC field 880 is entered in the same way as in the associated MARC fields.

| Following the removal of Tn datasets (records for undifferentiated names) from the Integrated Authority File (GND) on 15 June 2020, the field description has been cleaned up.

4. Further information and test data

The official Library of Congress documentation of the fields described can be found at https://www.loc.gov/marc/bibliographic/.

PDF and Excel versions of all documentation relating to the use of the MARC 21 format are available at https://www.dnb.de/EN/marc21.
In order to provide a brief overview of the changes made since publication of the previous description, all changes to the previous version are given on the first pages of the PDF versions in each case.
A sample file containing the notified changes can be found at https://data.dnb.de/testdat/ under file name "TitelTestRelease2020_03".

5. Unconfirmed advance information about the next but one release

Please note that this advance information on possible future changes is, by its nature, incomplete; changes may also be postponed to a later release or omitted entirely. Similarly, there may also be further major changes to the content. Which of the points are in fact implemented can be found in the corresponding release announcement which will be issued by November 2020.

Further information on the release cycle can be obtained from: https://wiki.dnb.de/x/wgcbBQ

For the next-but-one release in February 2021, the following changes are being planned:

Bibliographic data:

| Subject entries in fields 600 – 655 which are also delivered in field 689 will be marked accordingly so that the client can prevent them from being imported twice.
| The Open Access information and other licence information is being revised. The relevant proposal was accepted by the MARC Advisory Committee with minor amendments, and the changes in the MARC format were published as update no. 28[4]. The corresponding content will be transported to MARC fields 506 (Restrictions on Access Note) and 540 (Terms Governing Use and Reproduction). At the same time, the holdings data will be modified and MARC fields 506 (Restrictions on Access Note) and 845 (Terms Governing Use and Reproduction) will be introduced.
| Introduction of the new MARC field 884 (Description Conversion Information) to depict provenance information relating to the entire extant data record, e.g. regarding data imports to the DNB's bibliographic data. The field is derived from the MARC proposal „Description Conversion Information in the MARC 21 Bibliographic Format"[5].

If you have any questions please do not hesitate to mail us at metadatendienste@dnb.de or phone us on +49 69 1525-1630.

Kind regards,
on behalf of Deutsche Nationalbibliothek (German National Library)
Heike Eichenauer

*** Reading. Hearing. Understanding. German National Library ***

--

Heike Eichenauer
Deutsche Nationalbibliothek / German National Library
Digitale Dienste
Adickesallee 1
60322 Frankfurt
Telephone: +49 69 1525-1074
Fax: +49 69 1525-1636
mailto: h.eichenauer@dnb.de
https://www.dnb.de

---

[4] http://www.loc.gov/marc/mac/2019/2019-01.html
[5] http://www.loc.gov/marc/mac/2015/2015-03.html