

1500 - Code(s) für Sprache(n) des Textes und des Originals (nach DIN 2335 / ISO 639-2, 3 Kleinbuchstaben)

ILTIS-Handbuch, Titeldaten, Formatdokumentation, Feldbeschreibungen

Stand: 19.09.2017

PICA3 / Steuerzeichen	PICA+ / UF	W	Inhalt	MARC 21 / UF / Pos.
1500	010@	J	Code(s) für Sprache(n) des Textes und des Originals (nach DIN 2335 / ISO 639-2, 3 Kleinbuchstaben)	041
/1	\$a	J	Sprache des Textes	\$a
/3	\$c	J	Sprache des Originals	\$h
"\$E"	\$E	N	Kennzeichnung der Erfassungsart (Code)	
"\$H"	\$H	N	Herkunft (Code)	
"\$K"	\$K	N	Konfidenzwert maschinell ermittelter Sprachcode (1,000 - 0,000)	
"\$D"	\$D	N	Datum der automatisierten Erstellung (JJJJ-MM-TT)	

Indextyp/Schlüsseltyp	UF	Indexierungsroutine	ADI
COD/SCT	\$a	Sy (wortweise mit Sonderzeichen, nur für intellektuell vergebene Sprachcodes)	-
COD/SCO	\$c	Sy (wortweise mit Sonderzeichen)	-
SCM	\$a	Sy (wortweise mit Sonderzeichen, nur für maschinell vergebene Sprachcodes)	-
SCE	\$E	W, Ph	
SCH	\$H	W, Ph	
SCK	\$K	W, Ph	
SCD	\$D	Ph	

Verwendung

Das Feld ist in allen Satzarten zulässig. In der ZDB ist nur das Unterfeld \$a zulässig. In nach RDA katalogisierten und gekennzeichneten Datensätzen ist die "Sprache der Expression" ein Kernelement und daher obligatorisch zu erfassen.

Link zum ZDB-Format

http://www.zeitschriftendatenbank.de/fileadmin/user_upload/ZDB/pdf/zdbformat/1500.pdf

Für fortlaufende Ressourcen sind die Angaben des ZDB-Formats zu beachten.

Beschreibung des Feldinhaltes

In Feld 1500 ist das RDA-Element "Sprache der Expression" abgelegt. Es enthält Angaben zu Sprache(n) des Inhalts einer Ressource sowie bei Übersetzungen Angaben zu Sprache(n) des Originals in codierter Form gemäß DIN 2335/ISO 639-2/B.

Besteht die Ressource ausschließlich aus nicht nichtsprachlichem Inhalt, zum Beispiel Bilder oder Instrumentalmusik, wird ein entsprechender Code vergeben.

Sprache des Originals bei Übersetzungen meint die Originalsprache eines Werkes. Die Sprache, aus der die vorliegende Ressource übersetzt wurde, muss nicht immer die Originalsprache sein. Sofern das der Fall ist, wird ausschließlich die Originalsprache aufgeführt.

Für den Eintrag maschinell ermittelter Sprachcodes wurde das Feld um drei Unterfelder erweitert: \$E (Kennzeichnung der Erfassungsart), \$H (Herkunft), \$K (Konfidenzwert maschinell ermittelter Sprachcode) und \$D (Datum der automatisierten Erstellung). Die Unterfelder \$E, \$H und \$D enthalten Provenienz-Informationen.

Im Unterfeld \$E kann der Code m (= maschinell ermittelt) vergeben werden. Im Unterfeld \$H wird die Bezeichnung des Prozesses angegeben, aus dem die Daten in PICA/CBS übernommen wurden. Der Prozess zur automatischen Ermittlung der Sprachcodes mit Hilfe der Erschließungssoftware wird mit dem Code aep-lc (= durch die DNB vergebener Language Code aus der Averbis-Software) beschrieben. Das Unterfeld \$D zeigt das Datum der maschinellen Erstellung der Sprachcodes an. Das Unterfeld \$K zeigt den Konfidenzwert an, mit dem die Software die Sprache der Ressource ermittelt hat. Das Unterfeld \$K kann Werte zwischen 1,000 und 0,000 enthalten.

Ausführungsbestimmungen

Enthält die Ressource Text in maximal drei Sprachen, werden für alle Sprachen Sprachencodes erfasst, getrennt durch /1.

Weist die Ressource mehr als drei Textsprachen auf, wird ein spezifischer Code für die prägende Sprache vergeben, alle weiteren Sprachen werden durch den Code "mul" zusammengefasst.

Bei mehreren Sprachangaben ist die Kennung vor jedem Sprachencode zu erfassen. Die Codes für die Sprache der vorliegenden Ressource werden vor denen für die Sprache des Originals angegeben. Die Codes beziehen sich auf die Sprachen des Hauptteils einer Ressource, d. h. Vor- und Nachworte, Einleitungen, Anhänge, Register u.Ä. können vernachlässigt werden.

Bei Musikressourcen wird in der Regel auf die Angabe einer Originalsprache verzichtet.

Enthält die Ressource kein sprachgebundenes Material, ist der Code "zxx" zu erfassen. Für die Entscheidung, ob sprachgebundenes Material enthalten ist, werden Informationsquellen mit Titelangaben (z.B. Titelseite, Label, Kolophon) nicht berücksichtigt.

Bei der Erfassung der Sprache der Expression wird die Sprache von Begleitmaterial (z.B. Booklet) nicht berücksichtigt.

Die maschinell ermittelten ISO 639-2/B-Sprachcodes werden nur bei NP-Titeldaten (O-Sätze) eingetragen, sofern noch kein intellektuell vergebener oder durch Parallelabgleich ermittelter Sprachcode in Feld 1500 vorliegt. Bei späterer intellektueller oder durch Parallelabgleich erfolgreicher Sprachcodevergabe sollen maschinell ermittelte Sprachcodes gelöscht werden. Zunächst wird nur ein maschinell ermittelter Sprachcode eingetragen, mittelfristig auch einen zweiten und dritten Sprachcode geplant.

Der Sprachcode in UF \$a wird für den Index SCT nur dann indexiert, wenn 010@ \$E nicht vorhanden ist.

Codes

Liste der ISO-Sprachcodes: [Alphabetische Liste der Codes](#) oder [Alphabetische Liste der Sprachen](#).

Codes, die keiner bestimmten Sprache zugeordnet sind:

mul	zusammenfassender Code bei mehr als 3 Sprachen
mis	wird für bekannte Sprachen verwendet, für die es aber keinen Code gibt
und	wenn die Sprache nicht bekannt ist
zxx	kein linguistischer Inhalt

\$E	Kennzeichnung der Erfassungsart
m	maschinell gebildet
\$H	Herkunft
aep-lc	durch die DNB vergebener Language Code (LC) aus der Averbis-Software (AEP)

Beispiele

1500 /1ger

1500 /1ger/3eng

1500 /1ger/1eng/1fre

1500 /1ger/3und

1500 /1ger/3pol [Erläuterung: Die Sprache der Expression des zugrundeliegenden Werks ist polnisch; für die Übersetzung der Ressource wurde eine englische Fassung des Werks zugrunde gelegt --> Angabe in /3pol nicht /3eng/3pol]

1500 /1ger/1mul

1500 /1mis/3ger

1500 /1zxx [kein sprachgebundenes Material]

Formatbeispiel für Erfassung maschinell ermittelter Sprachcodes:

1500 /1gre \$Em \$H aep-lc \$K 0,554 \$D 2017-03-07

1500 /1eng \$Em \$H aep-lc \$K 0,511 \$D 2017-03-07
1500 /1fre \$Em \$H aep-lc \$K 0,478 \$D 2017-03-07

Altdaten / Datenpflege

<https://wiki.dnb.de/download/attachments/41157656/1500.pdf>