

Relevante Identifier

Inhalt:

Liste der relevanten Identifier

MAB2	MARC 21	Bezeichnung	RDF-Properties (Vorschläge zur Umsetzung)	Priorität
001	001	Identifikationsnummer des Datensatzes		
001_	001	_ = nicht definiert	<ul style="list-style-type: none"> cg:htn (HT-Nummer, regionaler hbz-Identifier) cg:bvn (regionaler BVB-Identifier) cg:ppn --> Allgemeines Prädikat für Identifier in Pica-Systemen) mit subproperties: <ul style="list-style-type: none"> cg:hebisppn (Regionaler HeBIS -Identifier) cg:gbvppn (Regionaler GBV-Identifier) cg:swbppn (Regionaler SWB-Identifier) Was ist mit dem KOBV? 	1
025	016 010 035	Überregionale Identifikationsnummer		
025a 025b 025c 025f 025i 025o 025z	016 7# \$a [...] \$2 "DE-101" 016 7# \$a [...] \$2 "Uk" 016 7# \$a [...] \$2 "ItFiC" 016 7# \$a [...] \$2 "FrPBN" 010 ## \$a [...] 035 ## \$a (OCoLC) [...] 016 7# \$a [...] \$2 "DE-600"	a = DNB b = BNB c = CasaliniLibri f = BNF l = LOC o = OCLC z = ZDB	<ul style="list-style-type: none"> cg:dnb http://open.vocab.org/terms/blid (Problem: Domain ist bibo:Document.) cg:casalinin cg:bnfn bibo:lccn (Problem: Domain ist bibo:Document.) bibo:oclcnum (Problem: Domain ist bibo:Document.) cg:zdbn 	
026	035	Regionale Identifikationsnummer		
026_	035 ## \$a (DE-599) [...]	_ = Erstkatalogisierung-Identifier	cg:eki (Es trägt sicher nicht zum Verständnis bei, dass es zum einen die EKI und zum anderen die EKI-URN gibt, die unterschiedliche Dinge identifizieren...)	1
540	020	Internationale Standardnummer (ISBN)		
540_ 540a	020 ## \$a [ISBN ohne Bindestriche] \$9 [ISBN mit Bindestrichen]	_ = ISBN formal nicht geprüft a = ISBN formal richtig	http://prismstandard.org/namespaces/basic/2.0/isbn <ul style="list-style-type: none"> bibo:isbn10 bibo:isbn13 	1
541	024	Internationale Standardnummer für Musikalien (ISMN)		
541_ 541a	024 2# \$a [ISMN ohne Bindestriche] \$9 [ISMN mit Bindestrichen]	_ = ISMN formal nicht geprüeft a = ISMN formal richtig		
542	022	Internationale Standardnummer für fortlaufende Sammelwerke (ISSN)		
542_ 542a	022 ## \$a	_ = ISSN formal nicht geprüeft a = ISSN formal richtig	<ul style="list-style-type: none"> http://rdvocab.info/Elements/issnOfSeries http://prismstandard.org/namespaces/basic/2.0/issn bibo:issn 	1
543	027	Internationale Standardnummer für Reports (ISRN)		
543_ 543a	027 ## \$a	_ = ISRN formal nicht geprüeft a = ISRN formal richtig		
551	028	Verlags-, Produktions- und Bestell-nummer von Musikalien und Tonträgern		
551a	028 52 \$a \$b	a = Verlags- und Firmenbestellnummer		
552	024	Digital Object Identifier		

552a 552b 552c	024 7# \$a [...] \$2 "doi" 024 7# \$a [...] \$2 "urn" 024 7# \$a [...] \$2 "handle"	a = DOI b = URN c = Handle	<ul style="list-style-type: none"> • bibo:doi • bibo:handle • http://prismstandard.org/namespaces/basic/2.0/doi • http://prismstandard.org/namespaces/basic/2.0/elsn • http://purl.org/spar/fabio/hasHandle Siehe http://purl.org/spar/fabio	
553	024	Artikelnummer		
553a	024 3# \$a	a = Internationale Artikelnummer (EAN)	bibo:gtin14	
554	502	Hochschulschriftennummer		
554_	502 ## \$o	_ = Hochschulschriftennummer	cg:unipubn (oder so)	
556	027 088	Reportnummer		
556a	027 ## \$a (wenn ISRN oder STRN) 088 ## \$a (wenn sonstige Reportnummer)	a = Reportnummer	cg:isrn (International Standard Technical Report Number)	
562	013	Patentnummer		
562a	013 ## \$a	a = Patentschrift		
578	026	Fingerprint		
578_	026 ## \$a \$b	_ = nicht definiert		
580	024	580 Sonstige Standardnummern		
580_	024 8# \$a	_ = VD16, VD17, VD18-Nummer; Hinweis: jeweils unterscheidbar an Präfix VD16, VD17, VD_18	<ul style="list-style-type: none"> • cg:vd16n • cg:vd17n • cg:vd18n 	
634	776	Internationale Standardbuchnummer (ISBN) der Sekundärform		
634_ 634a	776 1# \$z	_ = ISBN formal nicht geprüft a = ISBN formal falsch	s.o.	
635	776	Internationale Standardnummer für fortlaufende Sammelwerke (ISSN) der Sekundärform		
635_ 635a	776 1# \$x	_ = ISSN formal nicht geprüft a = ISSN formal richtig	s.o.	

Stand: 24.03.2011

Liste weiterer (für das Matching relevanter) Felder

Neben den oben genannten Felder für Identifikatoren gibt es noch weitere Felder, die wertvolle Hinweise darauf liefern, ob es sich um identische Datensätze handelt oder nicht. Diese Felder sind vor allem dann relevant, wenn der Datensatz nur einige wenige Identifikatoren enthält.

Die Konkordanzen zwischen den MAB-Felder und den MARC-Feldern wurden unter Berücksichtigung folgender [Liste](#) zusammengestellt.

MAB2	MARC 21	Bezeichnung
003	005 (Kontrollfeld)	Datum der letzten Bearbeitung
036	008 und 044	Ländercode
037	008	Sprachencode
038	041	Code für Herkunftssprache

050	007 und 008 und 090	Datenträger
051	008 und 090	Veröffentlichungsspezifische Angaben zu begrenzten Werken
052	006 oder 007 oder 008 oder 090	Veröffentlichungsspezifische Angaben zu fortlaufenden Sammelwerken
		Verfasser
100	100	Name der ersten Person in Ansetzungsform
102	100 oder 700	Identifikationsnummer des Personennamensatzes der ersten Person
		Körperschaften
200	110	Name der 1. Körperschaft in Ansetzungsform Körperschafts- oder Gebietskörperschaftsname als Eintragungselement
202	110 oder 710	Identifikationsnummer des Körperschaftsnamensatzes der ersten Körperschaft
400	259	Ausgabebezeichnung in normierter Form
403	250	Ausgabebezeichnung in Vorlageform
331	245	Hauptsachtitel
412	260	Verlagsname (falls mehrere bei DNB nur ersten auswerten)
410	260	Verlagsort
455	490	Bandangaben
425	260 oder 008	Erscheinungsjahr
433	300	Umfangsangabe

Stand: 01.07.2011

Anforderungen an die Datenlieferungen

Um die Datensätze analysieren zu können, benötigen wir Datenlieferungen, in denen die Felder der beiden Listen verarbeitet wurden. Lieferungen, in denen mehr als die notwendigen Felder exportiert wurden stellen kein Problem dar.

Ein Ergebnis des letzten Arbeitstreffen mit der AG KVA war die Übereinkunft, Datenlieferungen in MARC21 in UTF-8 im August zu liefern. Falls es keine Möglichkeit gibt MARC zu exportieren, wäre eine zweite Variante das MAB2-Format mit angegebenem Zeichennencoding (siehe unten).

Damit die Lieferungen der Daten problemlos verarbeitet werden können, benötigen wir eine gemeinsame Richtlinien bezüglich Namensgebung und Format.

Namenskonvention: VERBUNDKÜRZEL-DATUM-CHARACTERENCODING[-EXTRAINFORMATION].FORMAT

Beispiele:

- DNB-20110714-mab.mab2,
- DNB-20110714-utf8.marc21,
- DNB-20110714-utf8-Jahre1950bis1960.marc21,
- DNB-20110714-utf8-Teil1.marc21

Falls mehrere Dateien geliefert werden sollten, ist unsere Bitte, dass die Daten nicht in Unterordner verschachtelt werden.

Nachtrag: Es hat sich herausgestellt, dass die Datenlieferungen am besten zu verarbeiten sind, wenn sie im ".gz"-Format komprimiert wurden. Gleichzeitig erleichtert dies auch das Abholen der Daten von den FTP-Servern.

RDF-Properties für Identifikatoren

- Die *allgemeine Property für Identifikatoren*. <http://purl.org/dc/elements/1.1/identifier>
- Weitere Identifier-Properties aus der FaBiO.
 - <http://purl.org/spar/fabio/hasArticleIdentifier> ("A local identifier for an article within a periodical issue.")
 - <http://prismstandard.org/namespaces/basic/2.0/issueIdentifier> "An identifier used to identify a specific issue of a magazine, a journal or other resource.")
 - <http://purl.org/spar/fabio/hasNationalLibraryOfMedicineJournalId> ("An identifier for all those journals available from the National Library of Medicine repository.")
 - <http://purl.org/spar/fabio/hasPubMedCentralId> ("An identifier for all those bibliographic entities hosted by the PubMed Central repository.")
 - <http://purl.org/spar/fabio/hasPubMedId> ("An identifier for all those bibliographic records hosted by the PubMed repository.")

Hier sind bereits bekannte RDF-Properties für einige wenige Identifier-Typen angegeben. Die Erstellung eines [Identifier-Vokabulars](#) im culturegraph-Namensraum scheint allerdings sinnvoll vor dem Hintergrund, dass für die meisten Identifier bis jetzt keine Properties existieren.