Bibliographic Metadata Specifications

July 2023

See also: www.hathitrust.org/member-libraries/resources-for-librarians/contributor-toolkit/

HathiTrust requires bibliographic records sufficient to:

- Identify and collate records that each describe items exemplifying the same manifestation (e.g., show all records for a particular edition of a specific title).
- Identify and track records to their contributor.
- Facilitate record updates.
- Make an initial rights determination about each volume. Details about the information that is used to bibliographically determine rights attributes is available here: http://www.hathitrust.org/bib_rights_determination.

Records submitted should:

- be as complete as possible
- be valid per the MARC21 format
- in MARCXML format
- in utf8 encoding
- include OCLC number
- each describe a single item using a single 955 field with that item's description (one bibliographic record per item; multi-volume works should have the same record repeated for each volume)
- describe the print version of the digitized item

Records missing the following MARC fields or data elements will result in error or warning (see Metadata Submission Guide for a key to error messages):

- LDR
- 008
- OCLC number
- Local system record number
- 245$a (or 245$k, when applicable)
- Item identifier (barcode or other identifier)

The following conditions result in warning alerts and, as they have potentially significant impacts for resource discovery and access, should be addressed whenever possible:

- 008 values that are not 40 characters
- missing OCLC numbers
- multiple 245s

Please supply an explanation when you are unable to address these conditions.

Records should adhere to the following standards (MARC 21 Minimal Level Requirements are bolded):

Data Element	MARC		Required	Strongly Preferred	Specifications
MARC Leader	**LDR**		X		Must include 24 characters, not repeatable.
Bibliographic Format	**LDR/06**			X	*Value should be recognized Record type value from MARC Format for Bibliographic Data*
	LDR/07				*Value should be recognized Bibliographic level value from MARC Format for Bibliographic Data*
Contributor's local system record number	**001** (or contributor-identified location)		X		*If local system numbers are not in MARC 001, please specify location in coversheet; not repeatable*
Contributor's MARC org code	**003**			X	*Value should be one of recognized codes from the MARC Code List for Organizations:* *http://www.loc.gov/marc/organizations/*
Date and time of latest transaction	**005**			X	
Additional material characteristics	006			X	
Physical description	007			X	

Fixed-length data elements	**008**	X		Should include 40 characters and be encoded as a control field; not repeatable
Type of date/Publication status	**008/06**		X	
Publication start date	**008/07-10**			
Publication end date (if applicable)	**008/11-14**			
Publication place	**008/15-17**			
Material type descriptions **(as specified per material type)**	**008/18-34**			
Language	**008/35-37**			*Value should be one of recognized codes from the MARC Code List for Languages: http://www.loc.gov/marc/languages/*
LCCN	010		X	
ISBN	020		X	
ISSN	022		X	
OCLC #	035 $a		X	*Must be the master record number, **not** the institution record number. Please specify prefixes used in*

	(or contributor-identified location)			*coversheet. If OCLC numbers are not in 035$a, please specify location used in coversheet.*
Canceled/invalid local system number	035 $z		X	*When contributors migrate to a new ILS and then submit updated records to HathiTrust, they should retain their legacy system number and provide this in an 035$z, formatted as follows: (<MARC Org Code for the institution>)<prefix naming the legacy system><legacy system number>.*
Cataloging source	**040 $c**		X	*Value should be from recognized codes from the MARC Code List for Organizations: http://www.loc.gov/marc/organizations/*
Language	041		X	*Use for resources that include multiple languages. Value should be from recognized codes from the MARC Code List for Languages: http://www.loc.gov/marc/languages/*
Call/Classification number(s)	050		X	*Include at least one classification number from an established classification system recognized in the MARC 21 Format for Bibliographic Data, if nothing but a local call number is included please include documentation of the structure of the 09X fields and input conventions in your coversheet.*
	060			
	082			
	086			
	09X			
Main entry (NR)	1XX $a		X	

Uniform title (NR)		240 $a		X	
Title statement (NR)		**245 $ak**	X		Should not be repeated
Edition statement (NR)		250		X	
Publication statement		260 $abc		X	
Physical description		**300 $a**		X	
Series statement (if applicable)		490		X	
Added entries (if applicable)		70X-75X		X	
Series Added Entry – Uniform title (if applicable)		830		X	
Item description	Local item identifier (e.g., barcode) or ARK identifier (Internet Archive content)	955 $b (or contributor-identified location)	X		*Local item identifiers must be less than 32 characters and alpha characters must be case-insensitive.* *The ARK identifier will used as the local item identifier for volumes downloaded from the Internet Archive.* Not repeatable.
	Internet Archive identifier	955 $q (or contributor-identified location)	X		*If applicable - required for volumes downloaded from the Internet Archive.* If not in the 955$q, please specify location in the coversheet.

				Not repeatable.
	Enumeration and chronology	955 $v (or contributor-identified location)	X		*When applicable, where available. Enumeration/chronology information (date in particular) is required for rights determination of serial volumes.* *Preferred that enumeration and chronology statements are <200 characters.* Not repeatable. If not in the 955$v, please specify location in the coversheet.