Namespace Registration for Metadata Identifiers (META)

Namespace ID: META Requested of IANA.

Version: 1

Date: 2020-02-25

Registrant:

Name: Juha Hakala

E-mail: juha.hakala@helsinki.fi

Affiliation: Senior adviser, The National Library of Finland

Address: P.O.Box 15, 00014 Helsinki University, Finland.

Web URL: https://www.kansalliskirjasto.fi/en/

The National Library of Finland maintains Finnish translations of MARC 21 and Dublin Core metadata formats. Their elements have been identified with URLs, which the library would like to replace with URNs in order to benefit from URN resolution services.

Background:

According to ISO 5127, metadata is data about other data, documents or records that describes their content, context, structure, data format, provenance and/or rights attached to them. Metadata elements and their machine readable codes are specified in (cataloguing) formats. Libraries have been using a metadata format (MARC) since late 1960s, but since 90’s museums, archives and many other organizations have joined them.

As a part of general drive towards open linked data, organizations maintaining metadata formats have started to publish these specifications in the Web, but often in a form (e.g. a PDF document) which is not machine readable.

Even if element level documentation is accessible via HTTP URIs, there has been no coordination between metadata formats or even between translations of the same format on how these URIs have been assigned and what information has been made available.

Dublin Core community users Persistent URLs[[1]](#footnote-1) (PURLs) as identifiers[[2]](#footnote-2). DC metadata element Creator has a Persistent URL

<http://purl.org/dc/terms/creator>

in the /terms/ namespace, and PURL

<http://purl.org/dc/terms/1.1/creator>

in the /elements/1.1/ namespace. Different PIDs are necessary because term definitions in /terms/ namespace may differ from those in /elements/1.1/ namespace.

NOTE These PURLs identify creator-related metadata elements in the Dublin Core format, not the creators themselves. Creators have other identifiers, such as ORCiDs and ISNIs, which may be expressed as URIs (e.g. <http://www.isni.org/isni/0000000124292982>).

French translation of DCMI Metadata Terms[[3]](#footnote-3) uses the same PURLs, but they resolve to English texts. In order to access the French version, it is necessary to use URLs. URL for creator in French is <http://www.yoyodesign.org/doc/dcmi/dcmi-terms/index.html#terms-creator>. Translations to e.g. Czech, Japanese and Italian are similar in this respect: they rely also to the PURL of the English version.

MARC 21 identifier assignment policy differs from Dublin Core. Each translation has its own set of URIs. Library of Congress has assigned URL

<https://www.loc.gov/marc/bibliographic/bd245.html>

to the description of MARC 21 tag 245 (Title statement). URL of the Finnish description of this tag is

<https://marc21.kansalliskirjasto.fi/bib/20X-24X.htm#245>

and the Swedish description has URL

<http://www.kb.se/katalogisering/Formathandboken/Bibliografiska-formatet/210-249/#245>

There is a Web page listing all MARC 21 translations[[4]](#footnote-4), but URIs their elements are not interlinked.

NOTE Some XML-based metadata formats have XML namespaces:

AudioMD: <http://www.loc.gov/audioMD/>

MARCXML: <http://www.loc.gov/MARC21/slim>

PREMIS: <http://www.loc.gov/premis/rdf/v3/>

VideoMD: <http://www.loc.gov/videoMD/>

Unfortunately namespaces for AudioMD, MARCXML and VideoMD are not resolvable using e.g. HTTP, and functionality supported by the PREMIS namespace is insufficient - links to individual tags such as

|  |  |
| --- | --- |
|  | http://www.loc.gov/premis/rdf/v3/Copyright |

do not provide useful results, since they all take the client only to the beginning of the file containing all element descriptions. When implemented in this manner links to descriptions of metadata elements do not help human and machine users of metadata to understand the metadata provided.

Purpose:

Provide a uniform basis and a tool for identification of elements in metadata formats.

Functional improvements to the current URL-based identifiers of metadata elements.

Benefits:

1. Linking to alternative versions (e.g. human readable / machine readable) of element descriptions with one URN.

1. Automatic selection of the appropriate language (see below). For instance, a bibliographic record containing a link to the URI explaining MARC tag 245 in Finnish is only useful for users who understand Finnish, because the HTTP server holding these pages cannot redirect Swedish / English speaking users to MARC tag 245 pages on other HTTP servers that would be appropriate for them.
2. Having a URN namespace dedicated for metadata elements may improve co-ordination on how identifiers for metadata elements are created and what they resolve to. It may also encourage the organizations maintaining metadata formats to provide documentation separately for each metadata element and in a form more suitable to the Web than e.g. PDF.

If a URN:META identifier is assigned to the Library of Congress description of MARC 245 tag[[5]](#footnote-5), it is possible to direct a user to the descriptions of this tag in other languages[[6]](#footnote-6), depending on the language settings of the user’s Web client. This functionality can be implemented by supporting the HTTP Accept-Language header in the URN resolver. A prerequisite for this functionality is that all element URLs from translated versions are harvested to the resolver’s URN – URL mapping table. Since these URLs should be stable, keeping the links up to date is feasible.

If a network protocol used does not support language negotiation the required functionality may also be implemented with the URN R-component.

NOTE A URN should not be assigned if an element already has a (well-managed) PID.

Syntax:

The Namespace-Specific String (NSS) may consist of two parts:

 o a prefix consisting of a code identifying the metadata format and

 optional sub-namespace code(s) separated by a colon(s); and

* a string assigned by the format maintenance agency or a third party. Such strings may be constructed according to the local preferences as long as they are are aligned with the requirements of RFC 3986 and RFC 8141.

The following formal definition uses ABNF [[RFC5234](https://tools.ietf.org/html/rfc5234)].

 meta-nss = prefix "-" meta-string

 prefix = format-code \*( ":" subspc )

 ; The entire prefix is case insensitive.

 format-code = 1\*(ALPHA / DIGIT)

 ; As assigned by the National Library of Finland

 ; (identifies the metadata format to which the branch

 ; is delegated).

 subspc = 1\*(ALPHA / DIGIT)

 ; As assigned by the respective format maintenance agency.

 meta-string = path-rootless

 ; The "path-rootless" rule is defined in [RFC 3986](https://tools.ietf.org/html/rfc3986).

 ; Syntax requirements specified in [RFC 8141](https://tools.ietf.org/html/rfc8141) MUST be

 ; taken into account.

The following metadata format codes SHALL be used:

Descriptive metadata

Code Format(s) URL

BF BIBFRAME <http://www.loc.gov/bibframe/>

DC Dublin Core <https://www.dublincore.org/specifications/dublin-core/>

DDI DDI <https://ddialliance.org/explore-documentation>

EAD EAD <https://www.loc.gov/ead/>

IMARC INTERMARC <https://www.bnf.fr/fr/intermarc-bibliographique-de-diffusion>

LIDO LIDO <http://network.icom.museum/cidoc/working-groups/lido/>

MARC MARC 21 <https://www.loc.gov/marc/marcdocz.html>

MARCXML MARCXML <http://www.loc.gov/standards/marcxml/>

MIX MIX <http://www.loc.gov/standards/mix/>

MODS MODS <http://www.loc.gov/standards/mods/>

ONIX ONIX <https://www.editeur.org/8/ONIX/>

UNIMARC UNIMARC <https://www.ifla.org/unimarc>

Administrative metadata

Code Format URL

PREMIS PREMIS <http://www.loc.gov/standards/premis/>

TEXTMD textMD <https://www.loc.gov/standards/textMD/>

AUDIOMD audioMD <https://www.loc.gov/standards/amdvmd/>

VIDEOMD videoMD <https://www.loc.gov/standards/amdvmd/>

Cataloguing rules

Code Rules URL

ISBD ISBD <http://iflastandards.info/ns/isbd/elements/>

RDA RDA <https://www.rdaregistry.info>

One code may cover an entire family of formats (e.g. MARC Authority, Bibliographic and Holdings formats). Sub-namespaces may be used to differentiate formats within these format families if necessary.

National translations of metadata standards and cataloguing rules shall use the codes and URNs of the original specifications. Thus the Finnish translation of MARC 21 shall use the prefix MARC of MARC 21 and URNs assigned to the elements of the English version of the format. HTTP language negotiation will be used by the URN resolver to direct the client to the correct language version.

National variants of metadata formats (e.g. historical FINMARC format, which was based on equally outdated USMARC) should have their own format codes, since their tags may differ from the original ones. For instance, UKMARC tag 245 is not the same as USMARC tag 245, since in the former subtitle had its own tag, 248.

New codes will be added by the National Library of Finland on request.

The structure (if any) of the meta-string is determined by the authority for the prefix. Whereas the prefix is case insensitive, meta-strings MAY be case sensitive at the preference of the assigning authority; parsers therefore SHALL treat these as case sensitive, and any case mapping needed to introduce case insensitivity is the responsibility of the relevant resolution system.

A hyphen MAY be used as the delimiting character between the prefix and the meta-string. Within the meta-string, a hyphen MAY be used for separating different sections of the identifier from one another.

A colon SHOULD be used within the prefix only as a delimiting character between the format code and sub-namespace code(s), which splits the format specific namespace into smaller parts.

Maintenance agencies SHOULD NOT use in meta-strings characters requiring percent-encoding.

Rules for lexical equivalence:

Case insensitivity of the prefix must be taken into account when URN:META identifiers are analyzed.

META assignment:

National and international metadata format maintenance agencies may use URN META when they want to assign persistent identifiers for the metadata elements and tags of their formats, and provide URN-based access to machine or human readable descriptions about these metadata elements. For the time being these descriptions are unstructured text on Web pages.

The URN assigned to the element shall not change even if the description of the element is changed. URNs assigned to deleted elements shall not be re-used.

Metadata format maintenance agencies shall have procedures in place to make sure that the assigned URNs are unique and persistent. Since the number of metadata elements on formats is relatively low (at most a few hundreds) such procedures can be simple (e.g. URN can be based on the name of the element).

If a format has been translated to multiple languages, agencies maintaining the original version and its translations shall agree between themselves who will assign URNs to its elements, default value being the organization responsible of the original version of the format.

Security and Privacy:

URN:META identifiers do not have any known security or privacy issues.

Interoperability:

URN:META identifiers do not have any known interoperability related issues.

Resolution:

General

URNs in the URN:META namespace shall be resolvable.

URN:META namespace shall support URN to URL resolution service from the identifier to the page describing the identified metadata element, or another service which fulfills a relevant function within this context.

Resolution services may be maintained by the agencies maintaining the formats or a third party. If so, the agency maintaining the format and the agency shall agree on how the URN – URL mappings are maintained.

Locating the appropriate resolver

There is no automated mechanism for locating the correct URN resolver for a URN:META identifier. Once a format maintenance agency has decided to implement URNs, if makes independently the decision of which resolver shall support resolution of its URNs within the URN:META namespace.

Initially URN:META URNs will be represented as HTTP URIs. Eventually URN:META identifiers may become resolvable as such. Each sub-namespace may have its own resolver.

Example

Namespace URN:META:MARC: contains URNs which identify the tags of MARC 21 metadata formats (https://www.loc.gov/marc/) in various languages.

In this example, URNs are expressed as HTTP URIs which use (non-existent) URN resolver located http://example.com.

If the language setting of the client is English (and as a default) URNs [http://example.com/urn:meta:marc:<nss](http://example.com/urn%3Ameta%3Amarc%3A%3Cnss)> will be resolved at MARC 21 format specific pages at directories

<https://www.loc.gov/marc/bibliographic/>

<https://www.loc.gov/marc/authority/>

<https://www.loc.gov/marc/holdings/>

on the Library of Congress site.

EXAMPLE 1

URN of the MARC Bibliographic format tag 245 (Title Statement)

[http://example.com/urn:meta:marc:bd245](http://example.com/urn%3Ameta%3Amarc%3Abd245)

resolves as a default to URL

<https://www.loc.gov/marc/bibliographic/bd245.html>

EXAMPLE 2

URN of the MARC Authority format tag 100

[http://example.com/urn:meta:marc:ad100](http://example.com/urn%3Ameta%3Amarc%3Aad100)

resolves as a default to URL

<https://www.loc.gov/marc/authority/ad100.html>

unless the Accept-Language header setting is used to direct the user to a page in e.g. Finnish or Swedish.

Since pages for all tags in MARC 21 format for bibliographic data have been named in the Library of Congress Web site using the same syntax (file name is bdxxx.html, where xxx is the MARC tag), no URN – URL mapping table is required in the resolver. URNs can be mapped to URLs of their descriptive pages programmatically.

Using HTTP language negotiation the client may request this page in other languages.

EXAMPLE 3

If the language setting of the client is Finnish, URN

[http://example.com/urn:meta:marc:bd245](http://example.com/urn%3Ameta%3Amarc%3Abd245)

resolves to

<https://marc21.kansalliskirjasto.fi/bib/20X-24X.htm#245>

and if it is Swedish, to

<http://www.kb.se/katalogisering/Formathandboken/Bibliografiska-formatet/210-249/#245>

If the language setting is not supported, default (English) shall be used.

EXAMPLE 4

URN of the Dublin Core Terms namespace property Title

[http://example.com/urn:meta:dc:terms-title](http://example.com/urn%3Ameta%3Adc%3Aterms-title)

resolves as a default to

<http://purl.org/dc/terms/title>

and URN of the Elements 1.1. namespace property Title

[http://example.com/urn:meta:dc:elements-title](http://example.com/urn%3Ameta%3Adc%3Aelements-title)

resolves as a default to

|  |
| --- |
| <http://purl.org/dc/elements/1.1/title> |

Persistence

Persistence of URN:META resolution services depends on the persistence of metadata formats.

Element specific metadata about deprecated metadata formats such as USMARC, UKMARC or FINMARC may be hard to find and its form and content may not be suitable for URN resolution. See for instance UKMARC documentation at

[https://www.webarchive.org.uk/wayback/archive/20160107133726/http://www.bl.uk/bibliographic/ukmarc.html](https://www.webarchive.org.uk/wayback/archive/20160107133726/http%3A//www.bl.uk/bibliographic/ukmarc.html)

or FINMARC documentation at

<https://www.kiwi.fi/display/Marc21/FINMARC>

Additional documentation / information:

None

Revision Information:

This registration request is applicable to metadata formats and cataloguing rules listed above. Other formats and rules may be added in the future, with additional examples.

This request has been updated 2020-02-07. Examples including hyphen have been added, and resolver address <http://urn.fi> has been replaced by <http://example.com>.

1. <https://archive.org/services/purl/help> [↑](#footnote-ref-1)
2. DCMI Metadata Terms (in English) are available at

<https://www.dublincore.org/specifications/dublin-core/dcmi-terms/> [↑](#footnote-ref-2)
3. <http://www.yoyodesign.org/doc/dcmi/dcmi-terms/index.html> [↑](#footnote-ref-3)
4. <https://www.loc.gov/marc/translations.html> [↑](#footnote-ref-4)
5. Such URN can be assigned by the Library of Congress or a third party. [↑](#footnote-ref-5)
6. Provided that they have URLs, which is not always the case. [↑](#footnote-ref-6)