

◦ UNIMARC in RDF: Issues for UBC

UNIMARC in RDF Project members:

Gordon Dunsire
Consultant, Edinburgh, UK

Mirna Willer, PhD
Department of Information Sciences, University of Zadar

Predrag Perožić, PhD student
Department of Information Sciences, University of Zadar

Contents

- Background: IFLA models and ISBD standard in RDF
- UNIMARC Strategic Programme project: UNIMARC in RDF
- State of the art: UNIMARC namespaces and Mappings
- UBC, Granularity, Bottom-up methodology

Background: IFLA models and ISBD standard in RDF

- To publish bibliographic data as linked open data (LOD) in the Semantic Web
- To represent bibliographic standards in the form required by SW technology & standard: RDF
- RDF „sentence“: triple
 - Subject – Predicate – Object
 - Class – Property – Object
 - URI – URI – „Literal“ or URI [Subject - ...
 - Resource [URI] – ISBDproperty [URI] – „UNIMARC Manual“
 - Record ID-URI – **isbd/P1004** - „UNIMARC Manual“
- To define & publish namespaces (URIs) for the elements & value vocabularies of a model/standard

Background: IFLA models and ISBD standard in RDF

- IFLA: publication of models & ISBD in RDF (Resource Description Framework)
 - 2008: FRBR - FRAD – FRSAD, [FRBR Review Group](#)
 - 2009: ISBD, [ISBD Linked Data Study Group](#)
 - 2013: UNIMARC
- [Open Metadata Registry](#), management service

Welcome to The Registry!

This is the home page for the Open Metadata Registry (formerly the NSDL Registry).

The Metadata Registry provides services to developers and consumers of controlled vocabularies and is one of the first production deployments of the [RDF-based Semantic Web Community's Simple Knowledge Organization System \(SKOS\)](#)

How to begin

[Step-by-step illustrated instructions...](#)

Play in the [Metadata Registry Sandbox](#)

If you just want to experiment and play around a bit, go on over to the [Metadada Registry Sandbox](#). The code is exactly the same, but it's using a non-production database. (Please be careful not to kick any sand out of the box) NOTE: The [sandbox](#) has its own domain now. If you had bookmarked the [beta](#) site thinking that it was the sandbox (and we did say that), it will

Dublin Core Element Set	Dublin Core Metadata Initiative	2008-05-29	Jon Phipps
esL	owner1	2012-09-29	hello
feaa	Sabina Necula	2012-05-20	neculasabina
FOAF (Friend Of A Friend)	Metadata Management Associates	2008-08-05	Jon Phipps
FRAD model	Gordon Dunsire	2013-10-22	Gordon Dunsire (ifla)
FRBR Entities for RDA	Metadata Management Associates	2014-01-25	Jon Phipps
FRBRer model	Gordon Dunsire	2011-09-24	Gordon Dunsire (ifla)
FRSAD model	Gordon Dunsire	2011-09-24	Gordon Dunsire (ifla)
ISBD elements	Gordon Dunsire	2013-10-22	Gordon Dunsire (ifla)
ISBD elements (open)	Gordon Dunsire	2012-06-15	Gordon Dunsire (ifla)
ISIL	Carsten Klee	2013-08-05	Carsten Klee
JuniorResearcher	Necula Sabina	2011-07-14	sabinanecula
KACES	Sam Oh	2008-12-17	sam.oh
LCMPT Authority Data Elements	Jane	2012-12-28	P jane Smith
LOC American Cartoon Prints Collection	Mohamed Berray	2013-04-28	Mohamed Berray
manil	Grupo Wörterverzeichnis-master	2014-01-23	M luisa Diez
MARC 21 Elements 00X	Metadata Management Associates	2011-09-11	
MARC 21 Elements 0XX	Metadata Management Associates	2011-09-11	
MARC 21 Elements 1XX	Metadata Management Associates	2011-09-11	

E

Feedback

SPAR

Project UNIMARC in RDF

- Representation of UNIMARC in RDF
 - Funded for first year by PUC, 2013
 - Focus on UNIMARC Bibliographic format
 - IFLA Funding, 2014
 - Focus on UNIMARC Bibliographic format (cont.) & Authorities format
- To support production of datasets from UNIMARC catalogues
 - Used in Europe, North Africa, Russia, China, Japan
- To support linked data interoperability with related IFLA standards and beyond

UNIMARC namespaces

Element sets

- “Bibliographic” format has same focus as International Standard Bibliographic Description (ISBD)
 - The entity [bibliographic] Resource ~ Manifestation => RDF class
 - Attributes (UNIMARC „elements”) => RDF properties
- Lossless data requires finest level of granularity
 - UNIMARC coded subfields

Element sets:

tag + 1st ind. + 2nd ind. + subfield identifier

200ba = Title Proper (Title is not significant)

2001ba = Title Proper (Title is significant)

	A	B	C	D	E	F	G	H	I	
1	tag	tagCap	ind	ind1Cap	ind2	ind2Cap	su	subCap	definition	scopeNote
50	200	TITLE AND STATEMENT OF RESPONSIBILITY	0	Title is not significant	#	blank (not defined)	a	Title Proper	Relates a resource to the title proper (the chief title of the item, includes alternative title but excludes other title information (e.g. subtitles) and parallel titles).	
51	200	TITLE AND STATEMENT OF RESPONSIBILITY	0	Title is not significant	#	blank (not defined)	b	General Material Designation	Relates a resource to the general material designation.	Deprecated in ISBD.
52	200	TITLE AND STATEMENT OF RESPONSIBILITY	0	Title is not significant	#	blank (not defined)	c	Title Proper by Another Author	Relates a resource to the title proper by another author (the chief title of a work by an author other than the one responsible for the first title when they are both contained in the same item and when there is no collective title).	
53	200	TI S RI	1	Title is significant	#	blank (not defined)	a	Title Proper	The chief title of the item, includes alternative title but excludes other title information (e.g. subtitles) and parallel titles.	
54	200	TI S RI	1	Title is significant	#	blank (not defined)	b	General Material Designation	Text of general material designation	
55	200	TI S RI	1	Title is significant	#	blank (not defined)	c	Title Proper by Another Author	The chief title of a work by an author other than the one responsible for the first title when they are both contained in the same item and when there is no collective title.	
64	200	TITLE AND STATEMENT OF RESPONSIBILITY	1	Title is significant	#	blank (not defined)	d	Parallel Title Proper	The title proper in another language and/or script relating to a title proper appearing in a \$a or \$c subfield	
65	200	TITLE AND STATEMENT OF RESPONSIBILITY	1	Title is significant	#	blank (not defined)	e	Other Title Information	Subtitles and other title information that appear subordinate to the title proper on an item	
66	200	TITLE AND STATEMENT OF RESPONSIBILITY	1	Title is significant	#	blank (not defined)				

UNIMARC namespaces

Value vocabularies

- Coded information stored in tag block 1xx
 - Code lists specify notation, term, description, and scope
 - 100bba8 = Type of publication date code
- Represented as RDF/SKOS vocabularies
 - Italian and Portuguese translations – multilingual environment
 - Interoperability with vocabularies of other schema
- 12 published so far
 - For example: Target audience

Vocabulary: Show detail for UNIMARC: Target audience

- Detail**
- Concepts
- History
- Versions
- Maintainers

Detail	
Owner:	Gordon Dunsire
Name:	UNIMARC: Target audience
URL:	
Note:	Type of intended audience.
Community:	
Status:	Published
Language:	English

URI	
Base Domain:	http://iflastandards.info/ns/unimarc/terms/
Token:	tac

Browse...

- [Resource Owners](#)
- [Vocabularies](#)
- [Element Sets](#)
- [SPARQL](#)

Feedback

Vocabulary: Show detail for UNIMARC: Target audience

- Detail**
- Concepts
- History
- Versions
- Maintainers

?	Preferred Label ?	URI ?	Status	Updated ?	Actions
✓	adult, general	.../ns/unimarc/terms/tac#m	Published	2013-03-28 9:37	
✓	adult, serious	.../ns/unimarc/terms/tac#k	Published	2013-03-28 9:36	
✓	children, ages 9-14	.../ns/unimarc/terms/tac#d	Published	2013-03-28 9:35	
✓	juvenile, general	.../ns/unimarc/terms/tac#a	Published	2013-03-28 9:32	
✓	pre-primary, ages 0-5	.../ns/unimarc/terms/tac#b	Published	2013-03-28 9:33	
✓	primary, ages 5-10	.../ns/unimarc/terms/tac#c	Published	2013-03-28 9:34	
✓	unknown	.../ns/unimarc/terms/tac#u	Published	2013-03-28 9:38	
✓	young adult, ages 14-20	.../ns/unimarc/terms/tac#e	Published	2013-03-28 9:36	

8 results

<http://metadataregistry.org/concept/show/id/6346.html>

Vocabulary: UNIMARC: Target audience
Concepts: adult, general

Detail | **Properties** | **History**

Detail

Preferred Label: adult, general

Language: English

URI: <http://iflastandards.info/ns/unimarc/terms/tac#m> (RDF)

Top Concept?:

Status: Published

Properties

notation	m	English	Published
preferred label	adult, general	English	Published
preferred label	adulti in generale	Italian	Published
preferred label	adulto, grande público	Portuguese	Published

[List](#) | [Get RDF](#)

Browse...

- [Resource Owners](#)
- [Vocabularies](#)
- [Element Sets](#)
- [SPARQL](#)

Feedback

powered by symfony

125%

URI design templates

Element set granularity at subfield level with superstructure of fields (tags) and 2 qualifiers (indicators). Coded subfields refined by character position.

Value vocabulary
granularity at code level.
Hash URIs used if code list is small, or self-referential (“other”, etc.)

Tag	Ind 1	Ind 2	Subfield	CharPos	URI	Attribute
200	1	_ [blank]	a		2001_a	Title proper
100	_	_	a	17	100__a17	Target audience code 1

Vocabulary token	Code	URI	Vocabulary: Term
tac	m	tac#m	Target audience: adult, general

Mappings

- UNIMARC tags and subfields have corresponding ISBD “elements”
 - Now out-of-date (?) after publication of ISBD consolidated edition, 2011
 - Category of alignment relationship to be determined
 - Equivalent or broader/narrower
 - To be used as basis for sub-property mappings
- Mappings from UNIMARC to other vocabularies being developed: UNIMARC/ISBD

UNIMARC/Bibliographic			Align ment	ISBD		
Property	Label	Definition	A	Property	Label	Definition
2--	Description information block					
200_	Title and statement of responsibility	The title along with any other title information and statements of responsibility relating to the title including any of the preceding repeated in other languages (parallel titles, parallel statements of responsibility, etc.)	=	P1159	has title and statement of responsibility area	Relates a resource to a statement including the title proper, parallel titles, other title information, and statements of responsibility.
2000_				P1170	has title statement	Relates a resource to a statement of its title.
2001_				P1173	has title statement of series or multipart monographic resource	Relates a resource that is a series or multipart monographic resource to a statement of its title.
				P1172	has parallel title statement	Relates a resource to a statement of its parallel title
				P1012	has title	Relates a resource to a word or phrase, or a group of characters, usually appearing on the resource, that is the name of the resource or the work (or
200__a	has title proper	The chief title of the item.	= ◇	P1004	has title proper	Relates a resource to the title proper (the chief name of a resource, i.e. the title
200__b	General material designation	Text of general material designation				
200__c	Title proper by another author	The chief title of a work by an author other than the one responsible for the first title when they are both contained in the same item and when there is no collective title.	=	P1118	has title of individual work by different author	Relates a resource to the title of an individual work when the resource comprises two or more works with different statements of responsibility without a collective title.
200__d	Parallel title proper	The title proper in another language and/or script relating to a title proper appearing in a \$a or \$c subfield	=	P1005	has parallel title	Relates a resource to a title in another language or script than the title proper and presented as an equivalent of the title proper on the preferred source or sources of information.

UBC, Granularity

- Intellectual value of UNIMARC is preserved by a finest-grained semantic representation
- Data can always be dumbed-down to the level of coarseness (less granularity) required by applications
 - Processed with shared open maps
 - Including schema.org and dct!
 - And BIBFRAME too ...
- Data should be published without loss
 - For semantically rich applications
- Universal Bibliographic Control ~ Semantic Web

References

- Dunsire, Gordon, Mirna Willer. Standard library metadata models and structures for the Semantic Web. // Library hi tech news 28, 3(2011), 1-12
<http://www.emeraldinsight.com/fwd.htm?id=aob&ini=aob&doi=10.1108/07419051111145118>
- Dunsire, Gordon; Mirna Willer. UNIMARC and Linked Data. // IFLA Journal 37, 4(December 2011), 314-326, http://www.ifla.org/files/hq/publications/ifla-journal/ifla-journal-37-4_2011.pdf
- Dunsire, Gordon; Mirna Willer. Initiatives to make standard library metadata models and structures available to the Semantic Web. // World Library and Information Congress: 76th IFLA General Conference and Assembly. Meeting 149. Information Technology, Cataloguing, Classification and Indexing with Knowledge Management "Libraries and the Semantic Web", 10-15 August 2010, Gothenburg, Sweden. Dostupno na: <http://www.ifla.org/files/hq/papers/ifla76/149-dunsire-en.pdf>
- Willer, Mirna; Gordon Dunsire; Predrag Perožić. The UNIMARC in RDF project: namespaces and linked data. // IFLA World Library and Information Congress, 79th IFLA General Conference and Assembly, 17-23 August 2013, Singapore. Session 222 — Expanding MARC metadata services with linked open data — UNIMARC Core Activity (UCA), <http://library.ifla.org/156/1/222-willer-en.pdf>
- Dunsire, Gordon; Mirna Willer; Predrag Perožić. Representation of the UNIMARC bibliographic data format in Resource Description Framework. // DC-2013, Lisbon, Portugal, [International Conference on Dublin Core and Metadata Applications]: Linking to the Future, Lisbon, 2 - 6 September, 2013, <http://dcevents.dublincore.org/IntConf/dc-2013/paper/view/173/91>
- Willer, Mirna; Gordon Dunsire. Bibliographic information organization in the Semantic Web. Oxford: Chandos, 2013.