FROM UNIMARC BIBLIOGRAPHIC AND AUTHORITY RECORD TO LINKED OPEN DATA

Mix & match, or

Poor and popular vs. rich and lonely, or

Sub-property ladder?

Mirna Willer

Department of Information Sciences, University of Zadar, Croatia

mwiller@unizd.hr

&

Leonardo Jelenković

UNIBIS, d.o.o., Zagreb, Croatia

unibis@unibis.hr

Background (1)

➢ Research project aimed at publishing bibliographic and authority linked data of the Croatian Union Catalogue CROLIST which implements IFLA UNIMARC bibliographic and authority formats, 2012-2014

➢ Which ontology/vocabulary to use?
➢ What is the optimal methodology?

Which ontology/vocabulary to use?

- The most used vocabulary, that is, the one that has been applied to the highest number of records
- Vocabulary published by IFLA
- Map UNIMARC (not published in 2012) to one of the published vocabularies or wait for the PUC to publish UNIMARC namespaces
  - Map UNIMARC to which vocabulary ?
  - Choose one or several vocabularies for the same type of data, e.g., title?
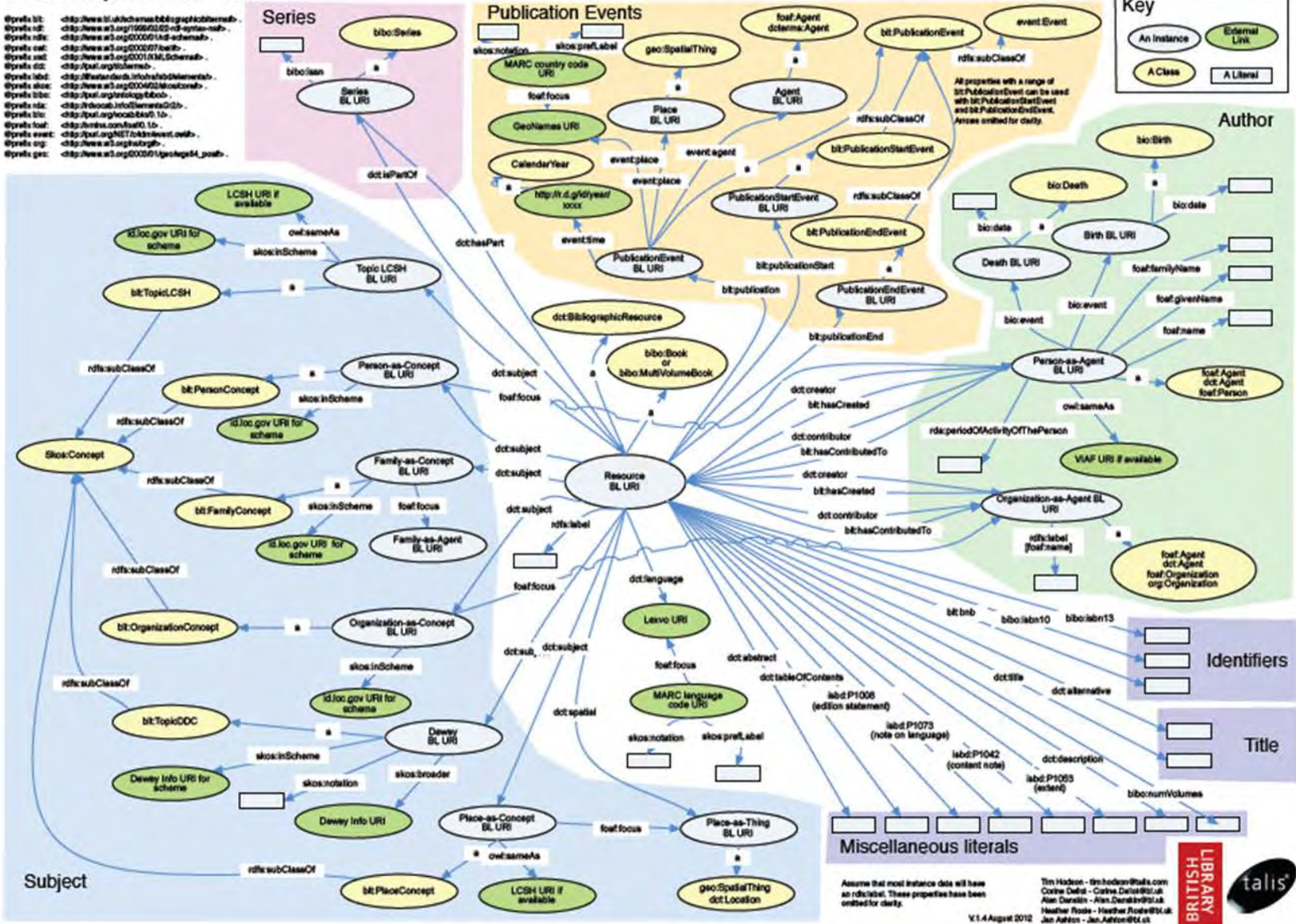
Methodology
1st approach

(UNI)MARC(21) map to

- a selection of elements of published from the vocabularies such as Dublin Core, ISBD, FRBR, MADS, RDA, Bibo, bnf-onto, Foaf, etc.

- No one single vocabulary meets the contextual and informational value of the source format

- *mix&match* method

- Examples: BL Data Model, BnF, BNE, DNB/Bibliographic, SUDOC, etc.

10/09/2014

IFLA Library Linked Data Satellite Meeting, Linked Data in Libraries: Lets' make it happen! Paris, 14 August 2014

4

http://www.bl.uk/bibliographic/pdfs/bldatamodelbook.pdf

Methodology
2nd approach

Use one (single) vocabulary

➢ Mapping from internal/local/national format to:

   ➢ a general one: Dublin Core, Schema.org, foaf, or

➢ Design one's own ontology: DNB gnd.org

➢ Publish local data in a local element set

   ➢ National Library of Scotland Digital Object Database elements

IFLA Library Linked Data Satellite Meeting, Linked Data in Libraries: Lets' make it happen! Paris, 14 August 2014

http://metadataregistry.org/schema/show/id/64.html

10/09/2014

IFLA Library Linked Data Satellite Meeting,
Linked Data in Libraries: Lets' make it
happen! Paris, 14 August 2014

7

http://metadataregistry.org/schema/show/id/64.html

open metadata registry
Supporting Metadata Interoperability

sign in / register | about

Search Vocabularies
Search Element Sets

Element Sets: Show detail for National Library of Scotland Digital Object Database elements

| Detail | Elements | History | Maintainers |

Label	Type	URI	Status	Updated	Last Updated by	Actions
has title	property	.../elements/tDescription-title	Published	2012-12-30 8:50	Gill Hamilton	
has description	property	.../tDescription-description	Published	2012-12-30 8:52	Gill Hamilton	
has parent	property	.../elements/hasResourceParent	New-Proposed	2012-07-15 6:16	Gill Hamilton	
has shelfmark	property	.../elements/hasResourceShelfmark	New-Proposed	2012-07-15 6:16	Gill Hamilton	
has Voyager bib id	property	.../hasResourceVoyagerBibId	New-Proposed	2012-07-15 6:17	Gill Hamilton	
has form or genre	property	.../dod/elements/hasFormOrGenre	Published	2012-12-30 8:04	Gill Hamilton	
has who	property	.../dod/elements/hasWho	Published	2012-12-30 8:12	Gill Hamilton	
has who depicted	property	.../elements/tWhoType-Depicted	New-Proposed	2013-01-24 2:32	Gill Hamilton	
is form or genre of	property	.../dod/elements/isFormOrGenreOf	Published	2013-01-24 2:33	Gill Hamilton	
has keyword	property	.../dod/elements/hasKeyword	Published	2012-12-30 7:55	Gill Hamilton	
is keyword of	property	.../dod/elements/isKeywordOf	Published	2012-12-30 7:57	Gill Hamilton	
is who depicted in	property	.../dod/elements/isWhoDepictedIn	New-Proposed	2012-08-11 5:57	Gill Hamilton	
has keyword label	property	.../dod/elements/tKeyword-keyword	Published	2012-12-30 8:36	Gill Hamilton	

Browse...
Resource Owners
Vocabularies
Element Sets
SPARQL

Feedback

125%

10/09/2014

IFLA Library Linked Data Satellite Meeting,
Linked Data in Libraries: Lets' make it
happen! Paris, 14 August 2014

8

Methodology
Implications of the 2nd approach

Publish one's data in one vocabulary only, whether a local or international:

➤ data expressed in one consistent vocabulary

➤ original contextual and informational value of data is preserved

➤ let other services reuse data according to their specific needs, and/or

➤ apply published namespaces maps:

    ➤ from a fine granularity element to a coarse(r) granularity element

➤ The *sub-property ladder* method(G. Dunsire)

Methodology
3rd approach

„Parralell" approach:

➢ Use UNIMARC namespaces (constructed according to the published methodology): basic IFLA standard vocabulary

    ➢ unimarcb, unimarca, ***and***

➢ map UNIMARC set of data elements to other standard and/or widely used

    ➢ bibliographic vocabularies: isbd, dc, bibo, foaf

    ➢ authority vocabularies: rdaRelGr2, gnd, gn, foaf, dc, edm, etc.

http://opak.crolib.hr/bibl/531114027640

```xml
<?xml version="1.0" encoding="UTF-8"?>
- <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:mime="http://purl.org/NET/mediatypes/"
  xmlns:bibo="http://purl.org/ontology/bibo/" xmlns:foaf="http://xmlns.com/foaf/0.1/" xmlns:dc="http://purl.org/dc/terms/"
  xmlns:isbd="http://iflastandards.info/ns/isbd/elements/" xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#" xmlns:rda="http://RDVocab.info/Elements"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
    - <rdf:Description rdf:about="http://opak.crolib.hr/bibl/531114027640">
        <rdfs:isDefinedBy rdf:resource="http://opak.crolib.hr/bibl/531114027"/>
        <foaf:page rdf:resource="http://opak.crolib.hr/page/531114027640"/>
        <isbd:P1032>531114027</isbd:P1032>
        <dc:type>monographic</dc:type>
        <dc:type>text</dc:type>
        <dc:language>hr</dc:language>
        <isbd:P1154>ISBN 978-953-14-1507-1</isbd:P1154>
        <dc:identifier rdf:resource="URN:ISBN 978-953-14-1507-1 "/>
        <bibo:isbn13>9789531415071</bibo:isbn13>
        <isbd:P1004>Čudnovate zgode šegrta Hlapića</isbd:P1004>
        <dc:title>Čudnovate zgode šegrta Hlapića</dc:title>
        <isbd:P1007>Ivana Brlić-Mažuranić</isbd:P1007>
        <isbd:P1016>Zagreb</isbd:P1016>
        <isbd:P1016 rdf:resource="http://dbpedia.org/resource/Zagreb"/>
        <isbd:P1017>Mozaik knjiga</isbd:P1017>
        <dc:publisher>Mozaik knjiga</dc:publisher>
        <isbd:P1018>2013</isbd:P1018>
        <dc:issued>2013</dc:issued>
        <isbd:P1022>129 str.</isbd:P1022>
        <dc:extent>129 str.</dc:extent>
        <isbd:P1023>ilustr. u bojama</isbd:P1023>
```

10/09/2014

IFLA Library Linked Data Satellite Meeting,
Linked Data in Libraries: Lets' make it
happen! Paris, 14 August 2014

11

http://opak.crolib.hr/auth/910306005

```xml
<?xml version="1.0" encoding="UTF-8"?>
- <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:unimb="http://iflastandars.info/ns/unimarc/unimarcb/elements/"
  xmlns:unima="http://iflastandars.info/ns/unimarc/unimarca/elements/"
  xmlns:rdaRelGr2="http://metadataregistry.org/uri/schema/RDARelationshipsGR2" xmlns:rdaGr2="http://rdvocab.info/ElementsGr2/"
  xmlns:rdaGr3="http://rdvocab.info/ElementsGr3/" xmlns:edm="http://europeana.eu/terms/"
  xmlns:gn="http://www.geonames.org/ontology#" xmlns:gnd="http://d-nb.info/gnd/" xmlns:bio="http://vocab.org/bio/0.1/"
  xmlns:foaf="http://xmlns.com/foaf/0.1/" xmlns:rel="http://purl.org/vocab/frbr/core#" xmlns:bibo="http://purl.org/ontology/bibo/"
  xmlns:dc="http://purl.org/dc/terms/" xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#" xmlns:rda="http://RDVocab.info/Elements"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
  - <rdf:Description rdf:about="http://opak.crolib.hr/auth/910306005182">
      <rdfs:isDefinedBy rdf:resource="http://opak.crolib.hr/auth/910306005"/>
      <foaf:page rdf:resource="http://opak.crolib.hr/paga/910306005182"/>
      <rdfs:label xml:lang="en">Description of Authority record: 910306005</rdfs:label>
      <unima:P000>00000nx a2200000n 45 </unima:P000>
      <rdf:type rdf:resource="http://purl.org/dc/terms/Agent"/>
      <rdf:type rdf:resource="http://xmlns.com/foaf/0.1/Person"/>
      <rdf:type rdf:resource="http://d-nb.info/gnd/NameOfThePerson"/>
      <unima:P001>910306005</unima:P001>
      <dc:identifier>910306005</dc:identifier>
      <gnd:gndidentifier>910306005</gnd:gndidentifier>
      <unima:P100bba>19910306ahrvb0191 ba</unima:P100bba>
      <dc:created rdf:datatype="http://www.w3.org/2001/XMLSchema#date">1991-03-06</dc:created>
      <dc:language rdf:resource="http://id.loc.gov/vocabulary/iso639-2/hrv"/>
      <unima:P152bba>PPIAK</unima:P152bba>
      <unima:P200b1a>Brlić-Mažuranić</unima:P200b1a>
      <foaf:familyName>Brlić-Mažuranić</foaf:familyName>
      <gnd:surname>Brlić-Mažuranić</gnd:surname>
```

10/09/2014

IFLA Library Linked Data Satellite Meeting,
Linked Data in Libraries: Lets' make it
happen! Paris, 14 August 2014

12

```xml
<unima:P200b1a>Brlić-Mažuranić</unima:P200b1a>
<foaf:familyName>Brlić-Mažuranić</foaf:familyName>
<gnd:surname>Brlić-Mažuranić</gnd:surname>
<unima:P200b1b>Ivana</unima:P200b1b>
<foaf:givenName>Ivana</foaf:givenName>
<gnd:forename>Ivana</gnd:forename>
<skos:prefLabel>Brlić-Mažuranić Ivana</skos:prefLabel>
<gnd:preferedName>Brlić-Mažuranić Ivana</gnd:preferedName>
<owl:sameAs rdf:resource="http://dbpedia.org/resource/Ivana_Brlić-Mažuranić"/>
<unima:P400b1a>Mažuranić</unima:P400b1a>
<foaf:familyName>Mažuranić</foaf:familyName>
<gnd:surname>Mažuranić</gnd:surname>
<unima:P400b1b>Ivana Brlić-</unima:P400b1b>
<foaf:givenName>Ivana Brlić-</foaf:givenName>
<gnd:forename>Ivana Brlić-</gnd:forename>
<skos:altLabel>Mažuranić Ivana Brlić-</skos:altLabel>
<gnd:variantName>Mažuranić Ivana Brlić-</gnd:variantName>
+ <gnd:variantName rdf:parseType="Resource">
<unima:P400b1a>Brlić Mažuranić</unima:P400b1a>
<foaf:familyName>Brlić Mažuranić</foaf:familyName>
<gnd:surname>Brlić Mažuranić</gnd:surname>
<unima:P400b1b>Ivana</unima:P400b1b>
<foaf:givenName>Ivana</foaf:givenName>
<gnd:forename>Ivana</gnd:forename>
<skos:altLabel>Brlić Mažuranić Ivana</skos:altLabel>
<gnd:variantName>Brlić Mažuranić Ivana</gnd:variantName>
- <gnd:variantName rdf:parseType="Resource">
    <gnd:surname>Brlić Mažuranić</gnd:surname>
    <gnd:forename>Ivana</gnd:forename>
```

🔍 125%

Methodology
What are the issues?

➢ Use of UNIMARC standard vocabularies

  ➢ *rich* in contextual informational value (no loss of data, fine granularity ensured) *and lonely* (in use)

&

➢ Use of general vocabularies such as DC, Bibo, Foaf, Scheme.org, etc., by library and non-library communities and users

  ➢ *poor* in contextual informational value ("dumbed-down" data) *and popular* (in use)

IFLA Library Linked Data Satellite Meeting, Linked Data in Libraries: Lets' make it happen! Paris, 14 August 2014

What are the benefits?

➢ Ensure consistency of data expressed by a particular vocabulary (*rich* or *poor*)

➢ Allow user services that "talk" – use a particular „language" – vocabulary to reuse the published data conforming to their requirements (semantic and/or technical)

➢ Retain the context of the data, and the informational value of UNIMARC vocabularies &

➢ Control and ensure the conversion of original UNIMARC vocabularies to other vocabularies

What are the challenges?

➢Recognising/Choosing a basic standard!

➢Relying on sustainable international standards (maintenance lifecycle)

➢Updating mappings of currently used vocabularies and adding new ones:

  ➢methodology and resources available (local & international expertise)

What are the challenges?

➢ Stable, persistent standard's namespaces:

  ➢ establish reliability, authenticity and accuracy of linked data

  ➢ improve interoperability / reuse and transparency of linked data

➢ Making RDF data available via a SPARQL [RDF Query Language] endpoint

➢ Representation of linked data: development of application profile(s)

References

➢ DCMI International Conference on Dublin Core and Metadata Applications, Lisbon, Portugal, 2013. Special Session: Data enrichment and transformation in the LOD context: Poor and popular vs. rich and lonely—Can't we achieve both?, http://dcevents.dublincore.org/IntConf/index/pages/view/rich-poor

➢ 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, G. Using the sub-property ladder, [blog] 2012, http://managemetadata.com/blog/2012/05/12/using-the-sub-property-ladder/

➢ Hillmann, D., G. Dunsire, J. Phipps. Maps and Gaps: Strategies for Vocabulary Design and Development. In *Proc. Int'l Conf. on Dublin Core and Metadata Applications 2013*, 82-89, http://dcevents.dublincore.org/IntConf/dc-2013/paper/view/185/80;

➢ Willer, M., G. Dunsire. Bibliographic information organization in the Semantic Web. Oxford: Chandos, 2013. Chap. 4, Publishing datasets as linked open data, pp. 199-248.

10/09/2014

IFLA Library Linked Data Satellite Meeting, Linked Data in Libraries: Lets' make it happen! Paris, 14 August 2014

18

UNIMARC zapis u RDF-u

```xml
- <rdf:RDF>
  - <rdf:Description rdf:about="http://opak.crolib.hr/bibl/9908170106(
      <isbd:P1032>990817010</isbd:P1032>
      <dc:identifier rdf:resource="http://opak.crolib.hr/bibl/99081701(
      <foaf:page rdf:resource="http://opak.crolib.hr/page/9908170106
      <dc:type>monographic</dc:type>
      <dc:type>text</dc:type>
      <dc:language>hr</dc:language>
      <isbd:P1154>ISBN 953-0-60425-4</isbd:P1154>
      <bibo:isbn>953-0-60425-4</bibo:isbn>
      <bibo:isbn10>9530604254</bibo:isbn10>
      <isbd:P1004>Čudnovate zgode šegrta Hlapića</isbd:P1004>
      <dc:title>Čudnovate zgode šegrta Hlapića</dc:title>
      <isbd:P1007>Ivana Brlić-Mažuranić</isbd:P1007>
      <isbd:P1008>16. izd</isbd:P1008>
      <dc:hasVersion>16. izd</dc:hasVersion>
      <isbd:P1016>Zagreb</isbd:P1016>
      <isbd:P1016 rdf:resource="http://dbpedia.org/resource/Zagreb"/>
      <isbd:P1017>Školska knjiga</isbd:P1017>
      <dc:publisher>Školska knjiga</dc:publisher>
      <isbd:P1018>1999.</isbd:P1018>
      <dc:issued>1999.</dc:issued>
      <isbd:P1022>130 str.</isbd:P1022>
      <dc:extent>130 str.</dc:extent>
      <isbd:P1023>ilustr.</isbd:P1023>
      <isbd:P1024>20 cm</isbd:P1024>
      <isbd:P1026 rdf:resource="http://opak.crolib.hr/bibl/930324110(
      <dc:title>Moja knjiga</dc:title>
      <isbd:P1017>Skolska knjiga</isbd:P1017>
      <dc:publisher>Školska knjiga</dc:publisher>
      <isbd:P1018>1999.</isbd:P1018>
      <dc:issued>1999.</dc:issued>
      <isbd:P1022>130 str.</isbd:P1022>
      <dc:extent>130 str.</dc:extent>
      <isbd:P1023>ilustr.</isbd:P1023>
      <isbd:P1024>20 cm</isbd:P1024>
      <isbd:P1026 rdf:resource="http://opak.crolib.hr/bibl/930324110632"/>
      <dc:title>Moja knjiga</dc:title>
      <isbd:P1029>Školska knjiga, Zagreb</isbd:P1029>
      <isbd:P1045>Str. 5-7: Putuje Hlapić svijetom / Marija Duš</isbd:P1045>
      <dc:description>Str. 5-7: Putuje Hlapić svijetom / Marija Duš</dc:description>
      <dc:isPartOf rdf:resource="http://opak.crolib.hr/bibl/930324110632"/>
      <dc:creator rdf:resource="http://opak.crolib.hr/auth/910306005"/>
      <dc:creator>Brlić-Mažuranić, Ivana</dc:creator>
      <dc:creator rdf:resource="http://dbpedia.org/resource/Ivana_Brlić-Mažuranić"/>
      <dc:creator rdf:resource="http://viaf.org/viaf/34584992"/>
      <dc:subject>UDC 821.163.42-93-31</dc:subject>
      <dc:contributor rdf:resource="http://opak.crolib.hr/auth/910328031"/>
      <dc:contributor>Duš, Marija</dc:contributor>
      <dc:contributor rdf:resource="http://viaf.org/viaf/68537339"/>
      <dc:contributor rdf:resource="http://opak.crolib.hr/auth/910328032"/>
      <dc:contributor>Antolčić, Ivan</dc:contributor>
      <dc:contributor rdf:resource="http://viaf.org/viaf/227427528"/>
      <dc:contributor>priredila Marija Duš</dc:contributor>
      <dc:relation rdf:resource="http://opak.crolib.hr/bibl/930324110632"/>
      <dc:coverage>HR, H</dc:coverage>
    </rdf:Description>
  </rdf:RDF>
```