

# MarcOnt – Integration Ontology for Bibliographic Description Formats

Sebastian R. Kruk, Marcin Synak, Kerstin Zimmermann

[sebastian.kruk@deri.org](mailto:sebastian.kruk@deri.org)  
[www.marcont.org](http://www.marcont.org)

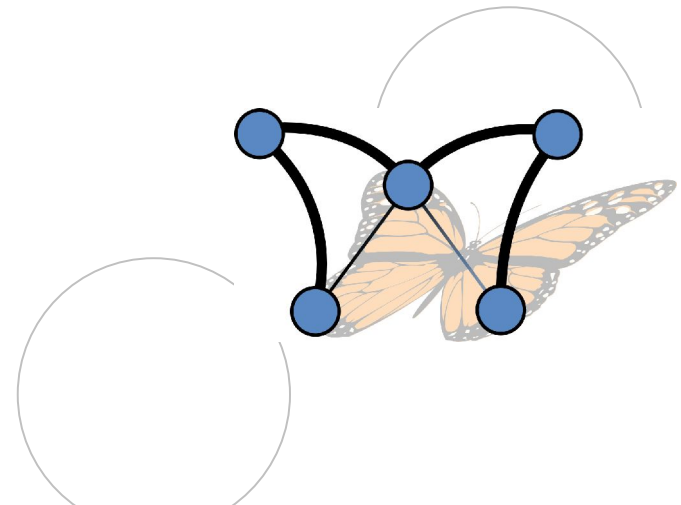
DC2005, Madrid  
13.09.2005

- Motivations
- MarcOnt Initiative Overview and Related Work
- Different types of Libraries
- Bibliographic Description Formats
- MarcOnt Ontology
- MarcOnt Mediation Services
- JeromeDL - proof of concept
- Future Work

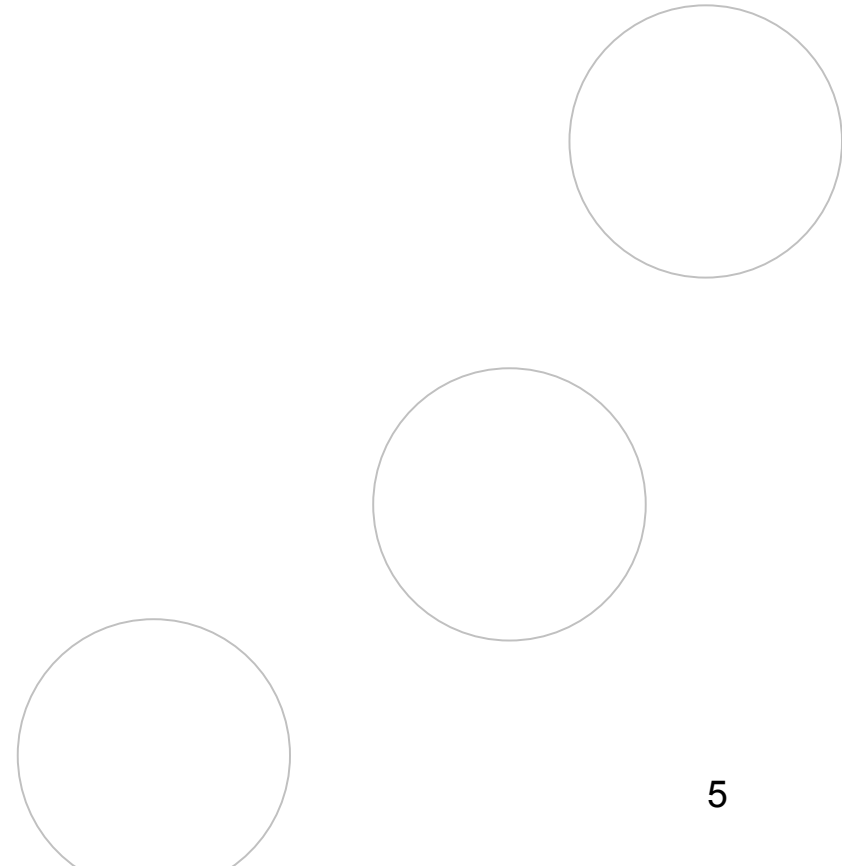
- Support for different kinds of bibliographic metadata, like: DublinCore, BibTeX and MARC21 at the same time.
- Utilize existing legacy bibliographic descriptions (like BibTeX or MARC21)
- Interlinking heterogeneous digital library networks.

# MarcOnt Initiative Overview and Related Work

- Goal: utilize the existing, legacy metadata in semantically enabled libraries
- Defines ontology that would cover concepts from MARC21, BibTeX and DublinCore
- Provides interoperability services to communicate with other entities using legacy metadata
- Similar solutions:
  - ABC ontology
  - GINF mediation framework



# Different Kind of Libraries (Evolution of Libraries)



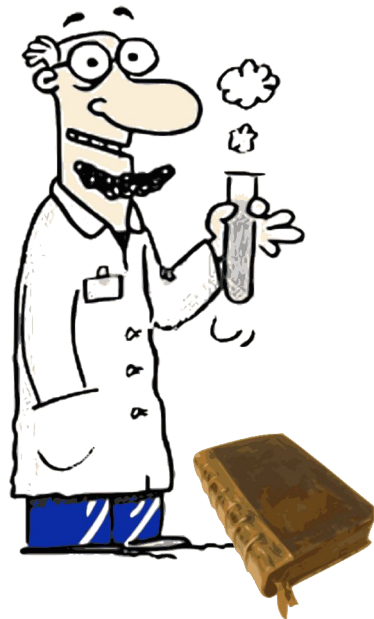
# Different Kind of Libraries (Evolution of Libraries)

- Classic Library



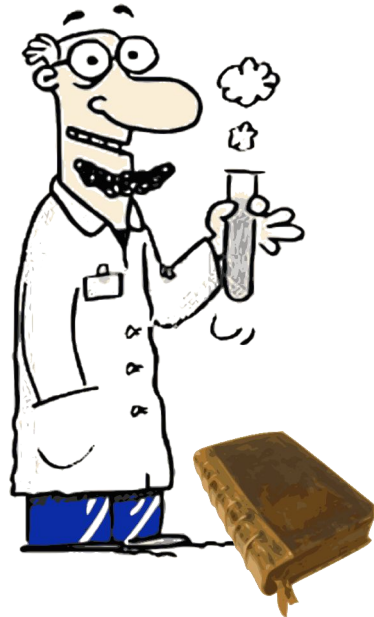
# Different Kind of Libraries (Evolution of Libraries)

- Classic Library
- Publishers' Library



# Different Kind of Libraries (Evolution of Libraries)

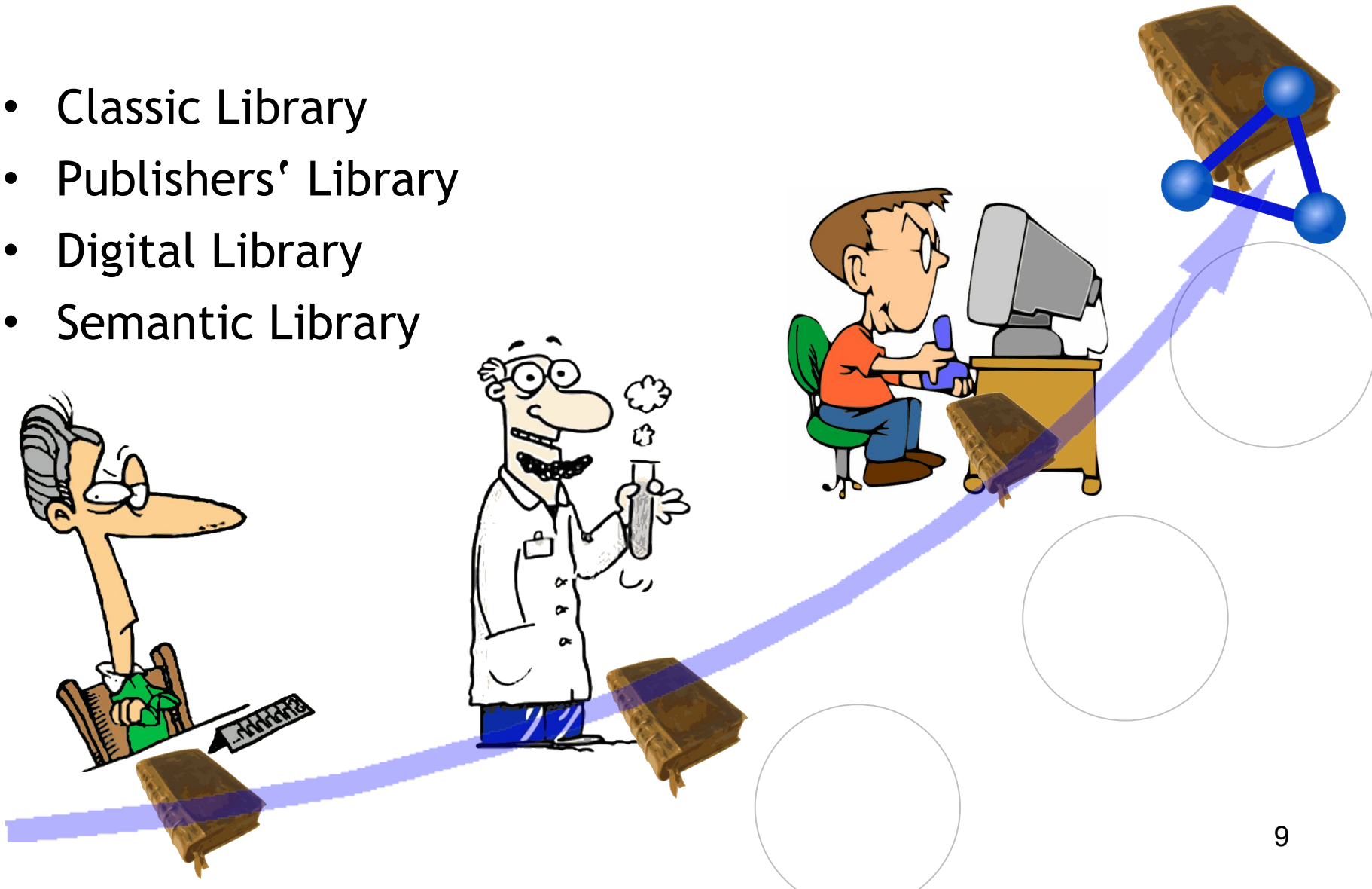
- Classic Library
- Publishers' Library
- Digital Library





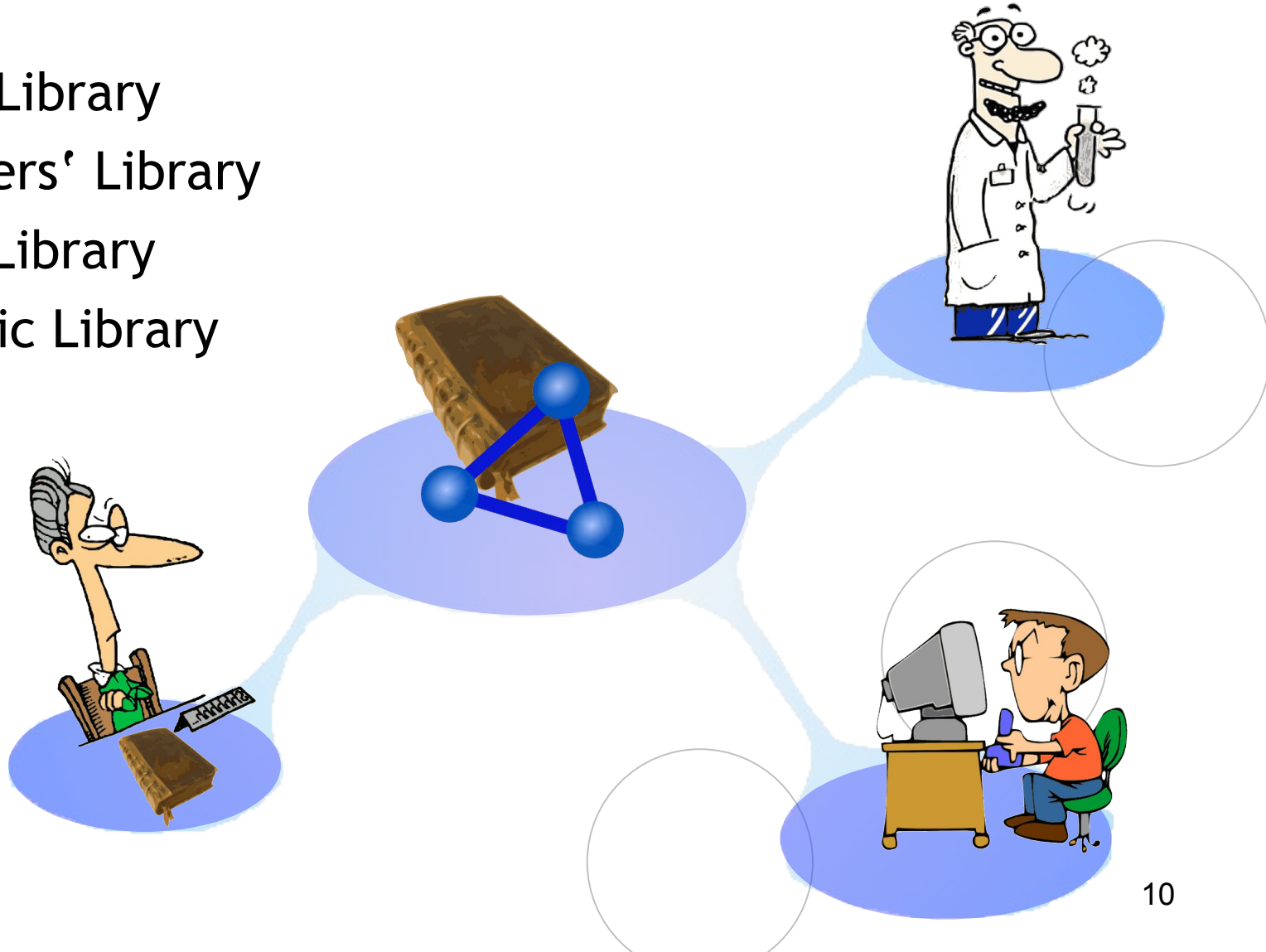
# Different Kind of Libraries (Evolution of Libraries)

- Classic Library
- Publishers' Library
- Digital Library
- Semantic Library



# Different Kind of Libraries (Evolution of Libraries)

- Classic Library
- Publishers' Library
- Digital Library
- Semantic Library



- **MARC21:**
  - numerical fields 001-887
  - 8 major types of material
  - 7 major types of records
  - number of various classifications
- **BibTeX:**
  - variable elements
  - 12 types of publications
- **Dublin Core:**
  - 15 meta elements
  - 12 types of categories

# Different Types

Books  
Continuing resources  
Computer files  
Maps  
Music  
Sound recordings  
(non music)  
Visual materials  
Mixed materials

Language (textual)  
material  
Manuscript (textual)  
language material  
Computer file  
Cartographic material  
Manuscript cartographic  
material  
Notated music  
Manuscript music

**MARC21**

String  
Book  
InBook  
InCollection  
Proceedings  
InProceedings  
Article  
MasterThesis  
PhDThesis  
TechReport  
Manual  
Misc

**BibTeX**

Collection  
Dataset  
Event  
Image  
MovingImage  
StillImage  
Sound  
Text  
Interactive resource  
Physical object  
Service  
Software

**DublinCore**

01450cas 922004331i  
 450000100...**019c19329999**gw qr | p |  
 ||||**0** |**0ger** | a0044-2992  
 9a200412140219bVLOADc20040407152  
 5dvkulc200310071018dvbjc200303101  
 205dkopumky200209211341zVLOAD  
 aGD U/MPcGD U/MPdGD U/MFdGD  
 U/KKsdWR O/EJ0 ager1 aZ.  
 Kunstgesch. 0aZeitschrift für  
 Kunstgeschichte00aZeitschrift für  
 Kunstgeschichte.18aZfK aMünchen  
 ;aBerlin :bDeutscher  
 Kunstverlag,c1932-. c26-29 cm.  
 aKwart.0 a1 Bd. (Juni 1932)-. aOpis  
 na podst.: LCC. aW 1932  
 założycielami czasopisma byli  
 Wilhelm Waetzoldt i Ernst Gall....

```
@InProceedings { jeromedexa2005,
  author = "Sebastian Ryszard Kruk and ... ",
  title = "{JeromeDL - Adding Semantic ...}",
  booktitle = "{In Proceedings to DEXA 2005}",
  year = 2005}
```

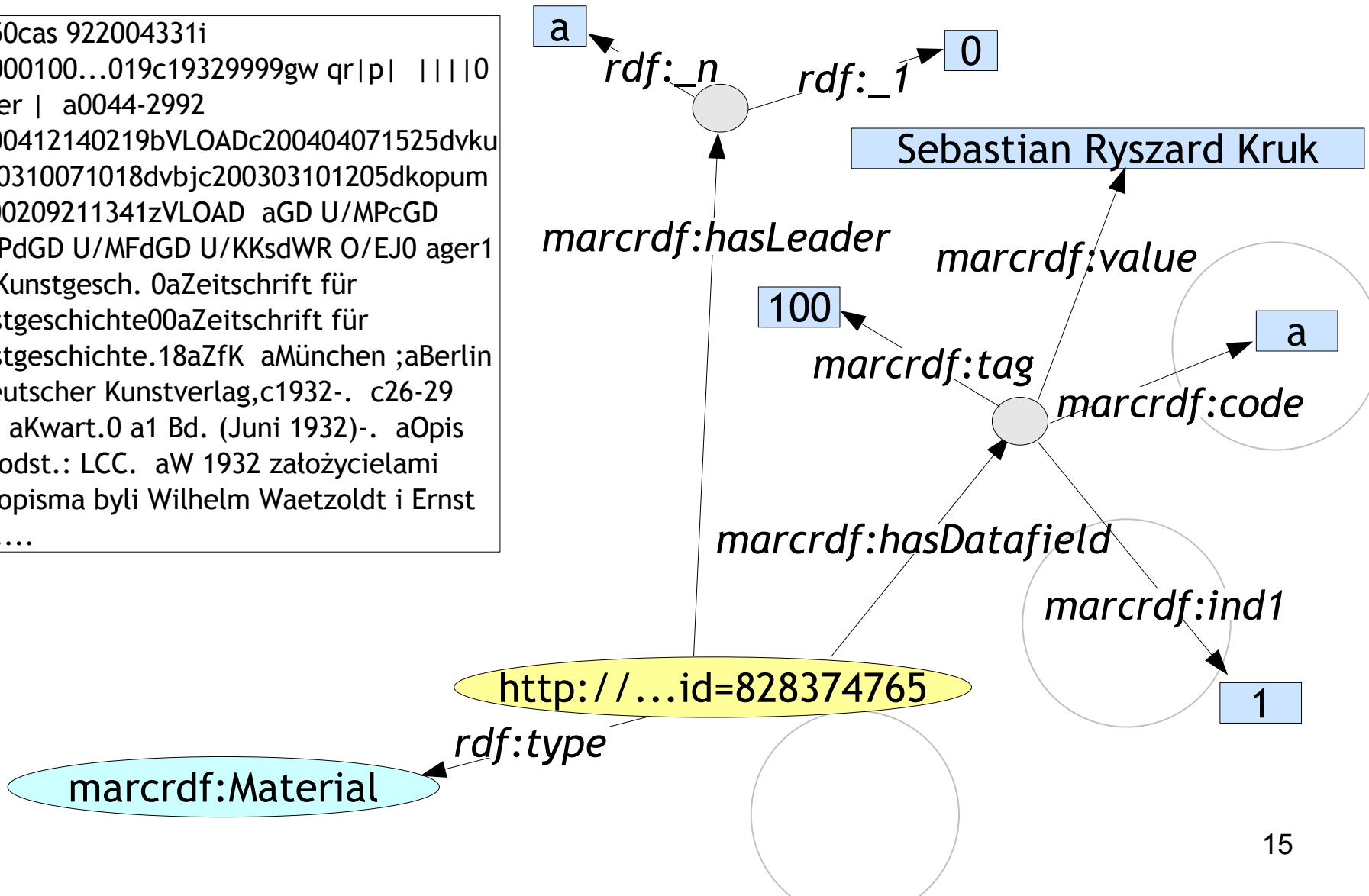
```
<?xml version="1.0" encoding="UTF-8" ?>
<Description>
  <dc:title>JeromeDL - Adding Semantic
  Web Technologies to Digital
  Libraries</dc:title>
  <dc:creator>Sebastian Ryszard
  Kruk</dc:creator>
  <dc:type> ... </dc:type>
</Description>
```

# Interoperability challenges

- providing access in a joined query language
- unified data representation – RDF
- unified vocabulary – ontology

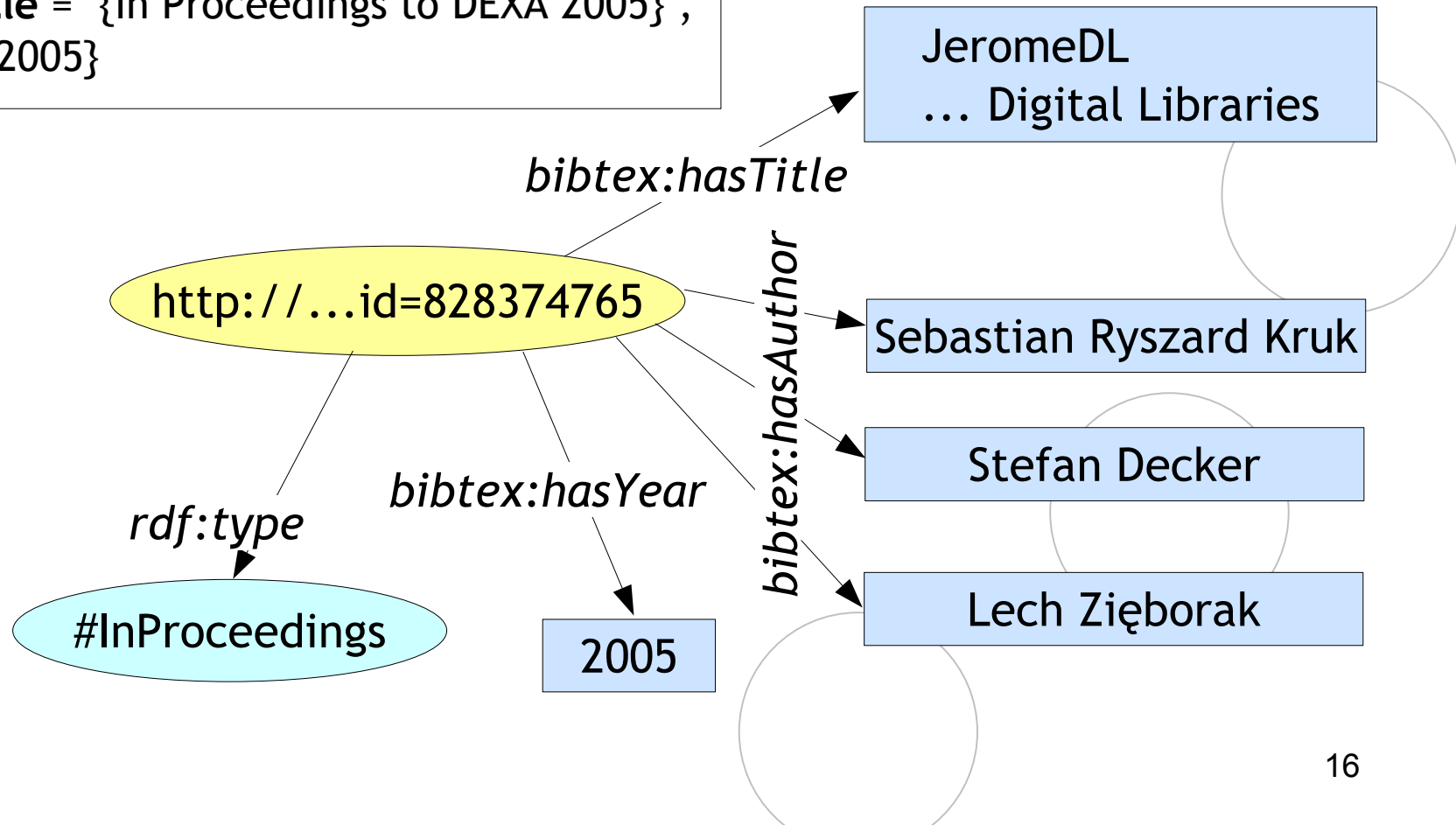
# Bibliographic Description - MARC21

01450cas 922004331i  
 450000100...019c19329999gw qr|p| ||||0  
 |0ger | a0044-2992  
 9a200412140219bVLOADc200404071525dvku  
 lc200310071018dvbjc200303101205dkopum  
 ky200209211341zVLOAD aGD U/MPcGD  
 U/MPdGD U/MFdGD U/KKsdWR O/EJ0 ager1  
 aZ. Kunstgesch. 0aZeitschrift für  
 Kunstgeschichte00aZeitschrift für  
 Kunstgeschichte.18aZfK aMünchen ;aBerlin  
 :bDeutscher Kunstverlag,c1932-. c26-29  
 cm. aKwart.0 a1 Bd. (Juni 1932)-. aOpis  
 na podst.: LCC. aW 1932 założycielami  
 czasopisma byli Wilhelm Waetzoldt i Ernst  
 Gall....



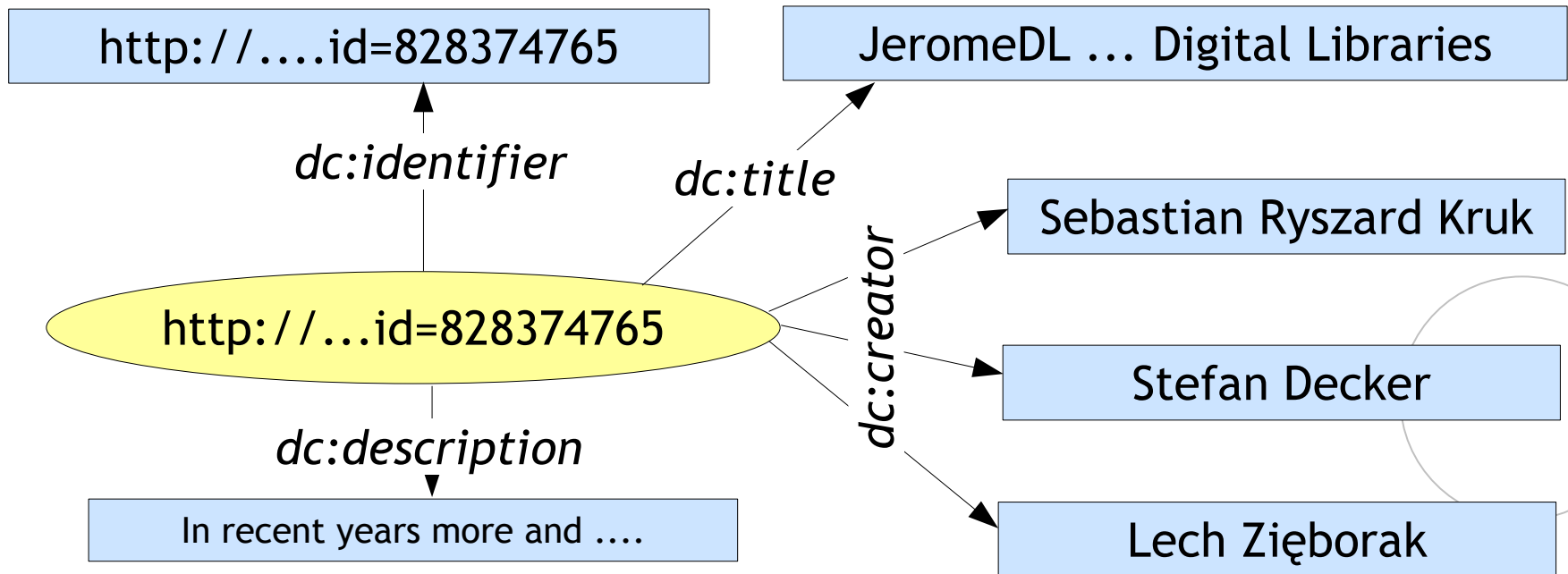
# Bibliographic Description - BibTeX

```
@InProceedings { jeromedexa2005,
  author = "Sebastian Ryszard Kruk and ... ",
  title = "{JeromeDL - Adding Semantic ...}",
  booktitle = "{In Proceedings to DEXA 2005}",
  year = 2005}
```



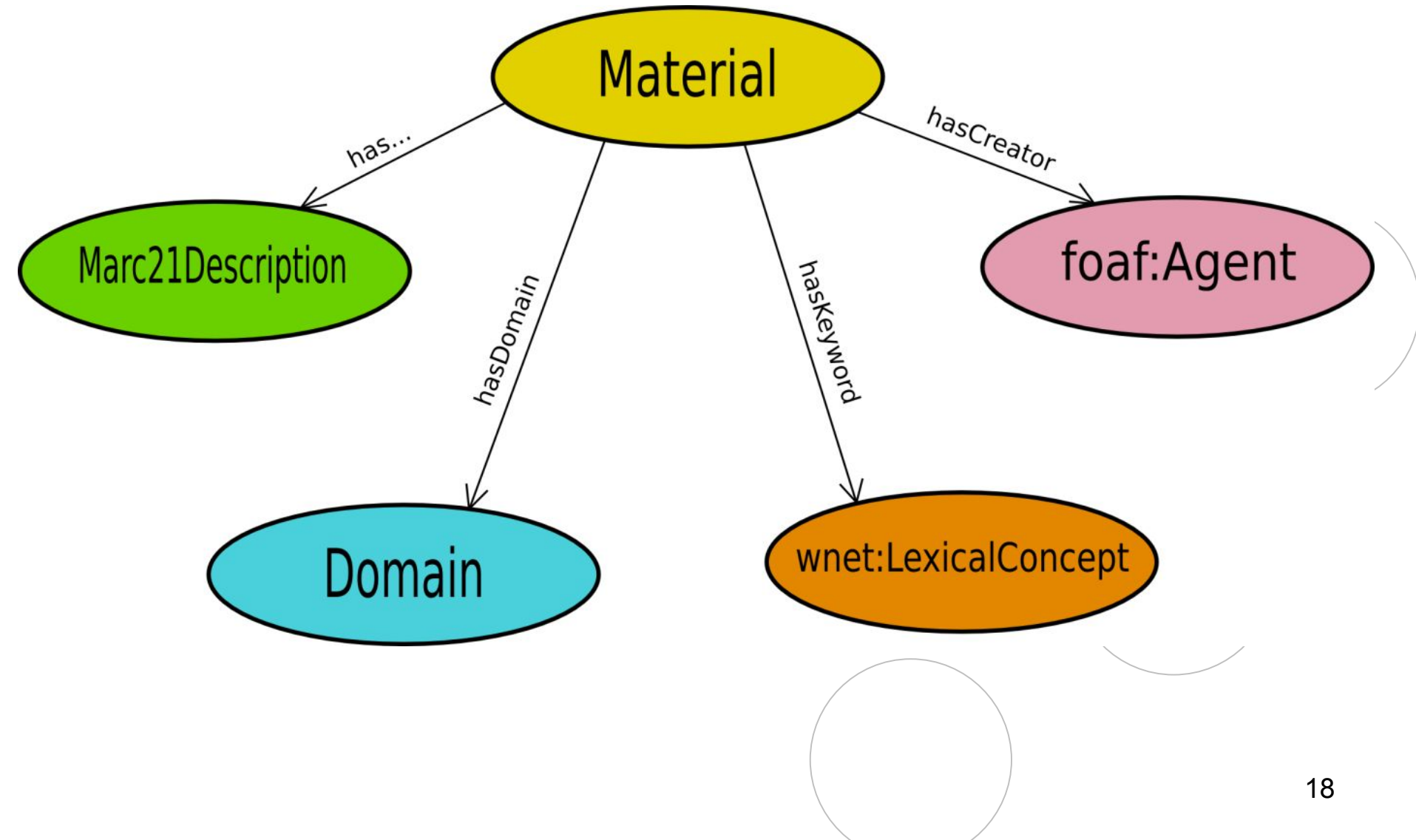


# Bibliographic Descriptions in RDF – Dublin Core



```
<?xml version="1.0" encoding="UTF-8" ?>
<Description>
  <dc:title>JeromeDL - Adding Semantic Web Technologies to Digital
  Libraries</dc:title>
  <dc:creator>Sebastian Ryszard Kruk</dc:creator>
  <dc:type> ... </dc:type>
</Description>
```

# Common vocabulary - MarcOnt Ontology



01450cas922004331i450000100..  
.019c19329999gw qr|p| |||0 |  
0gera004429929a200412140219  
bVLOADc20040407

@InProceedings {jeromedexa2005,  
author = "Sebastian Kruk and ...",  
title = "{JeromeDL - ...}",  
year = 2005}

<Description>  
<dc:title>MarcOnt ...</dc:title>  
<dc:creator> Kruk</dc:creator>  
</Description>



# MarcOnt Mediation Services

## MarcOnt ontology

*RDF Storage*

MARC-RDF  
input/output  
adaptor

MARC-RDF

MARC-XML

MARC21

BibTeX-RDF  
input/output  
adaptor

BibTeX-RDF

BibTeXML

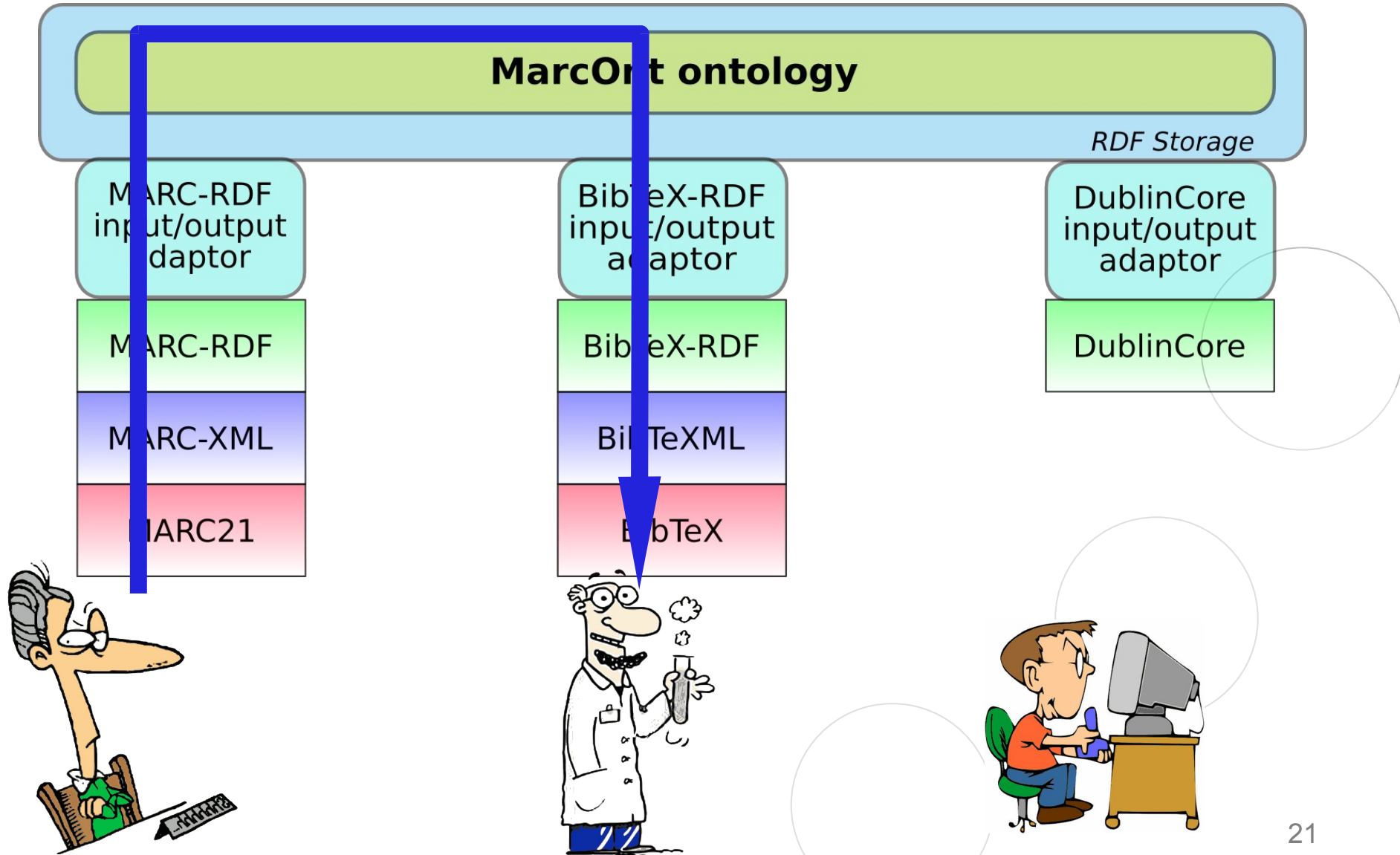
BibTeX

DublinCore  
input/output  
adaptor

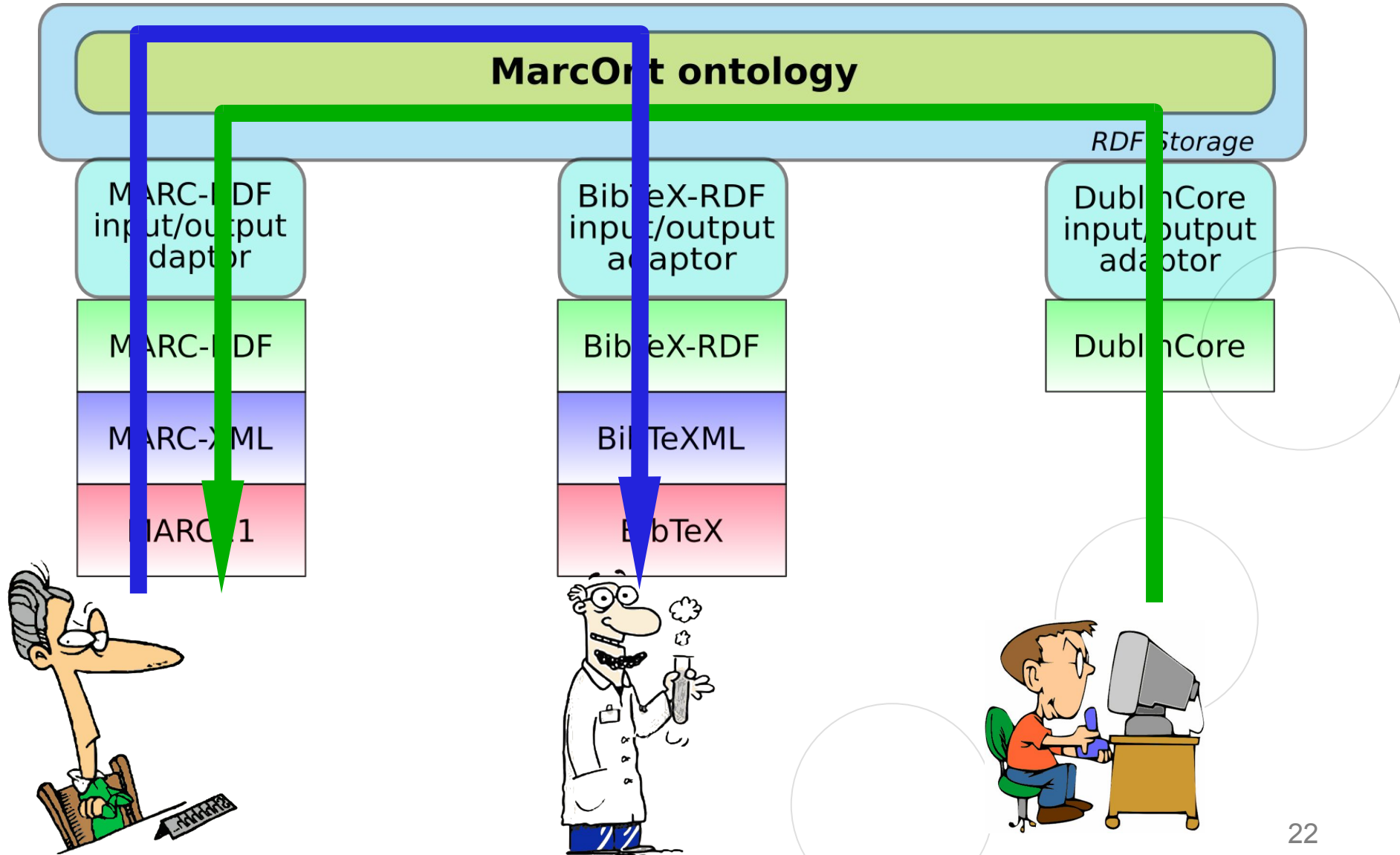
DublinCore



# MarcOnt Mediation Services

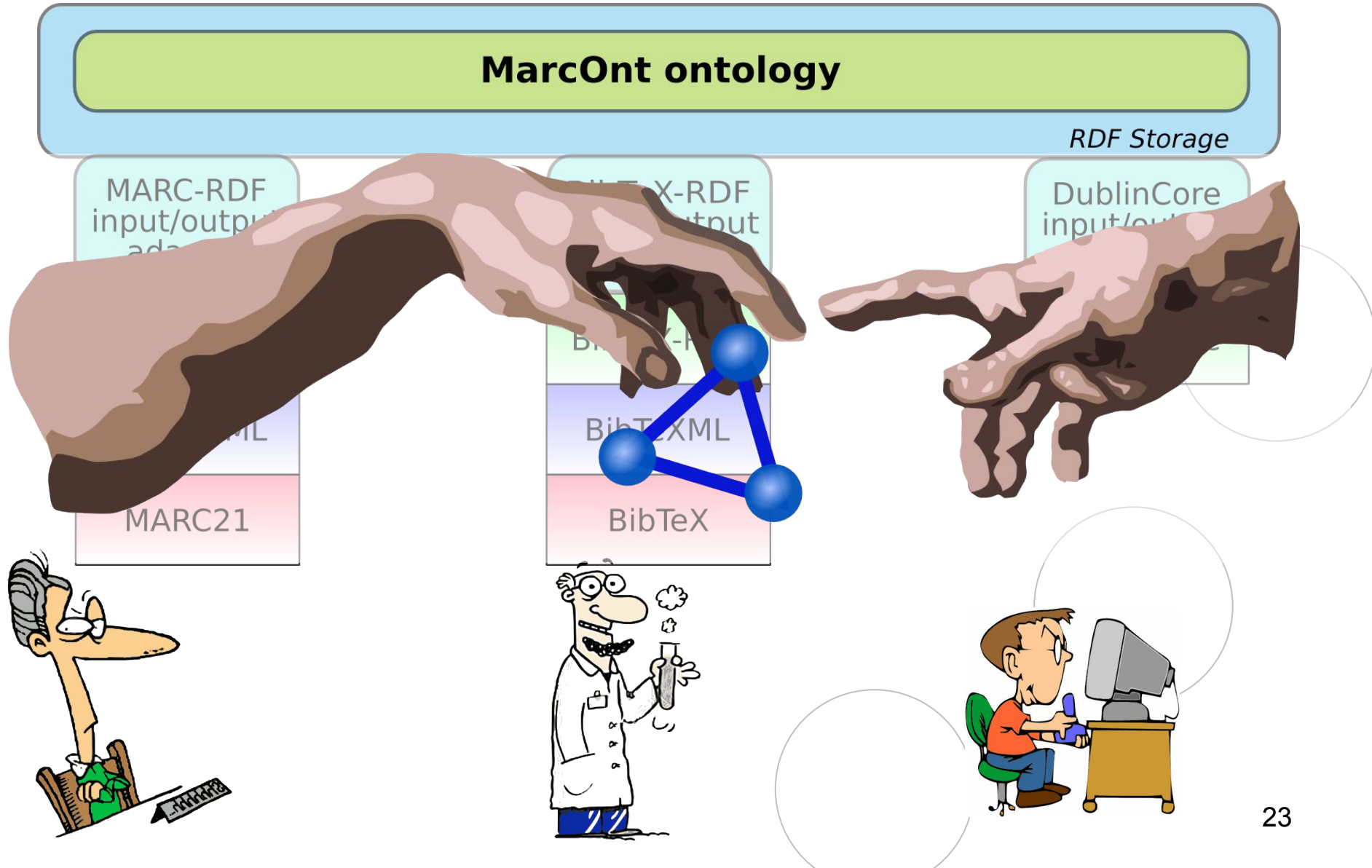


# MarcOnt Mediation Services

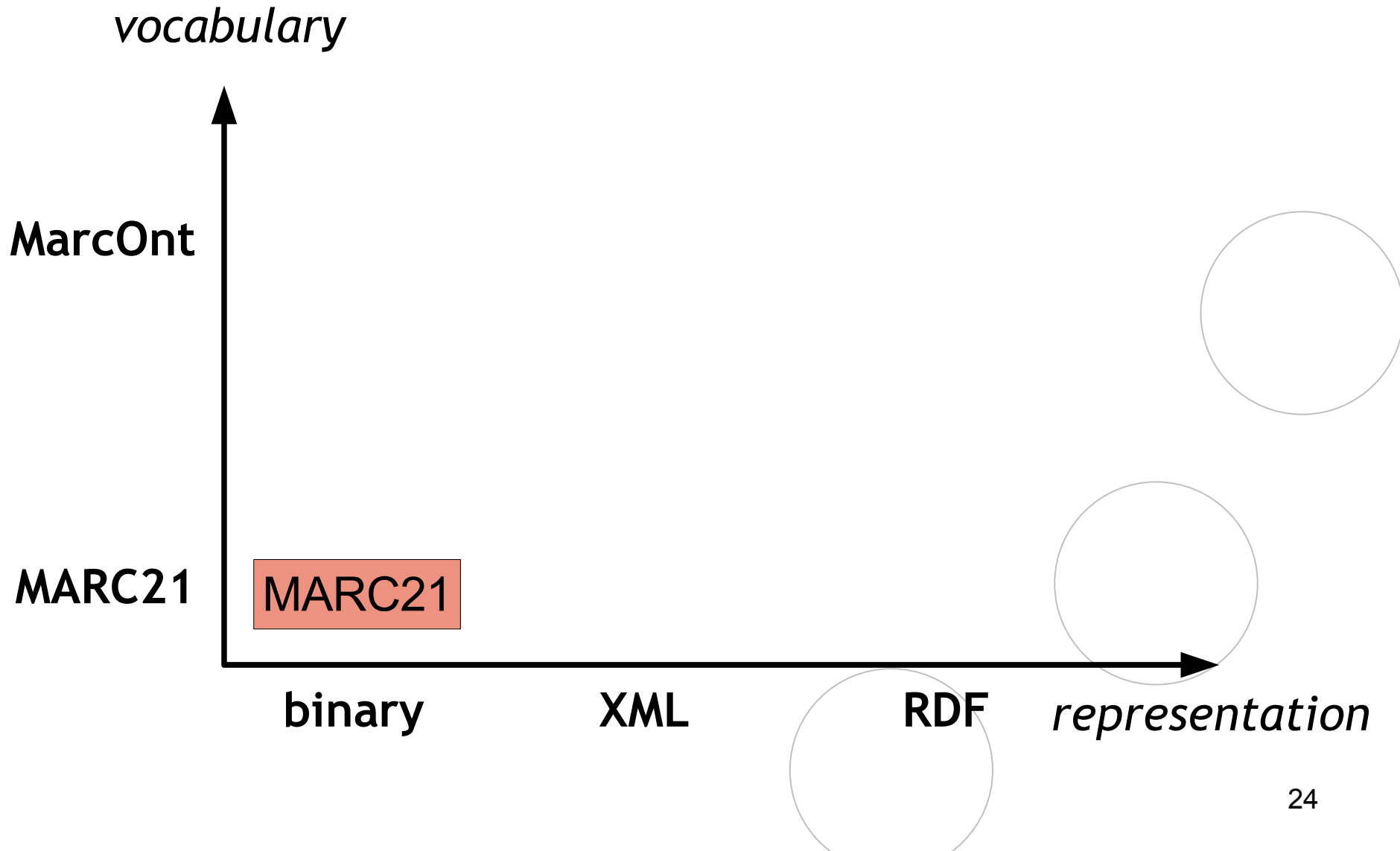




# MarcOnt Mediation Services

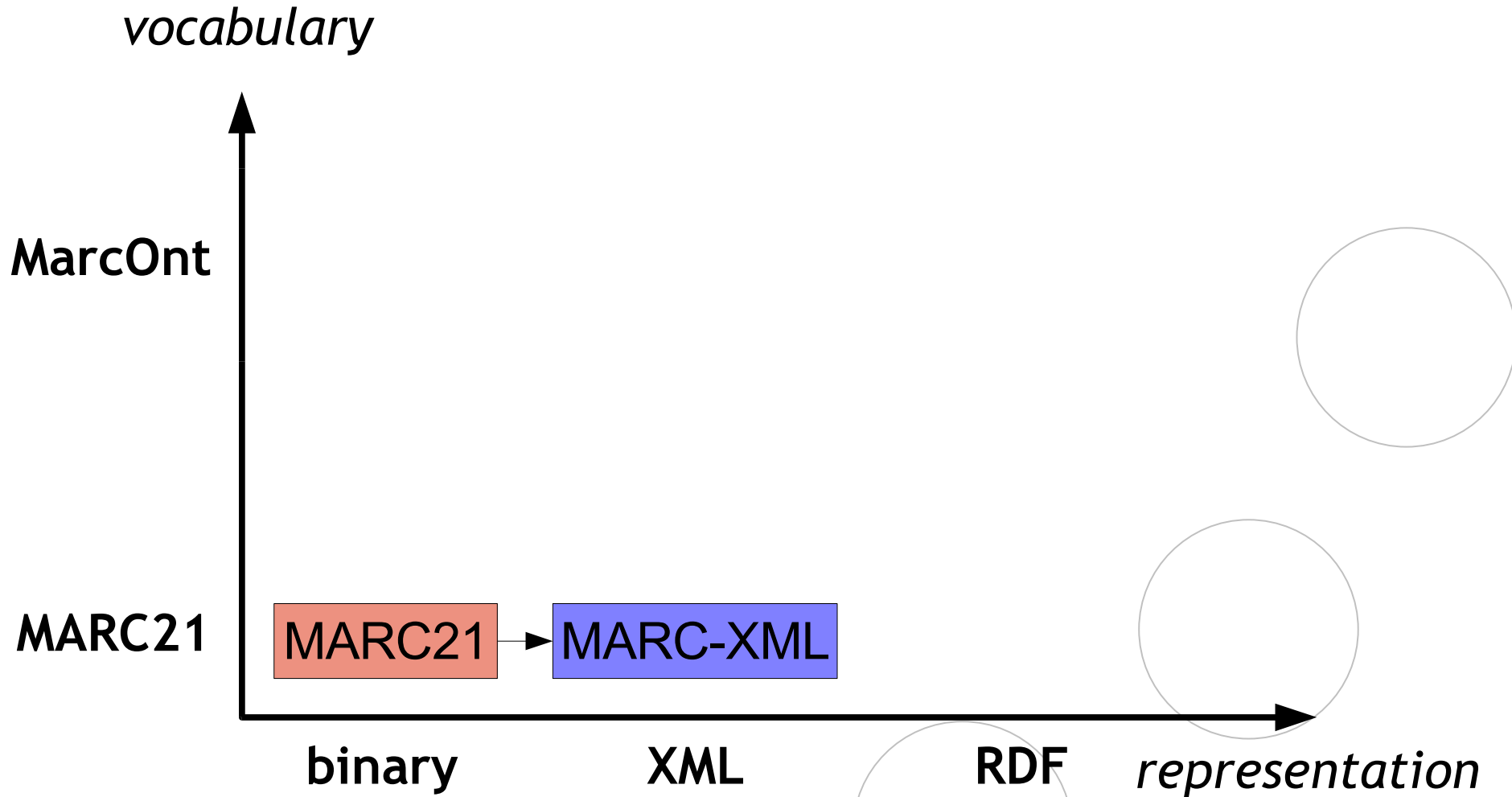


# From Legacy Description to MarcOnt Ontology

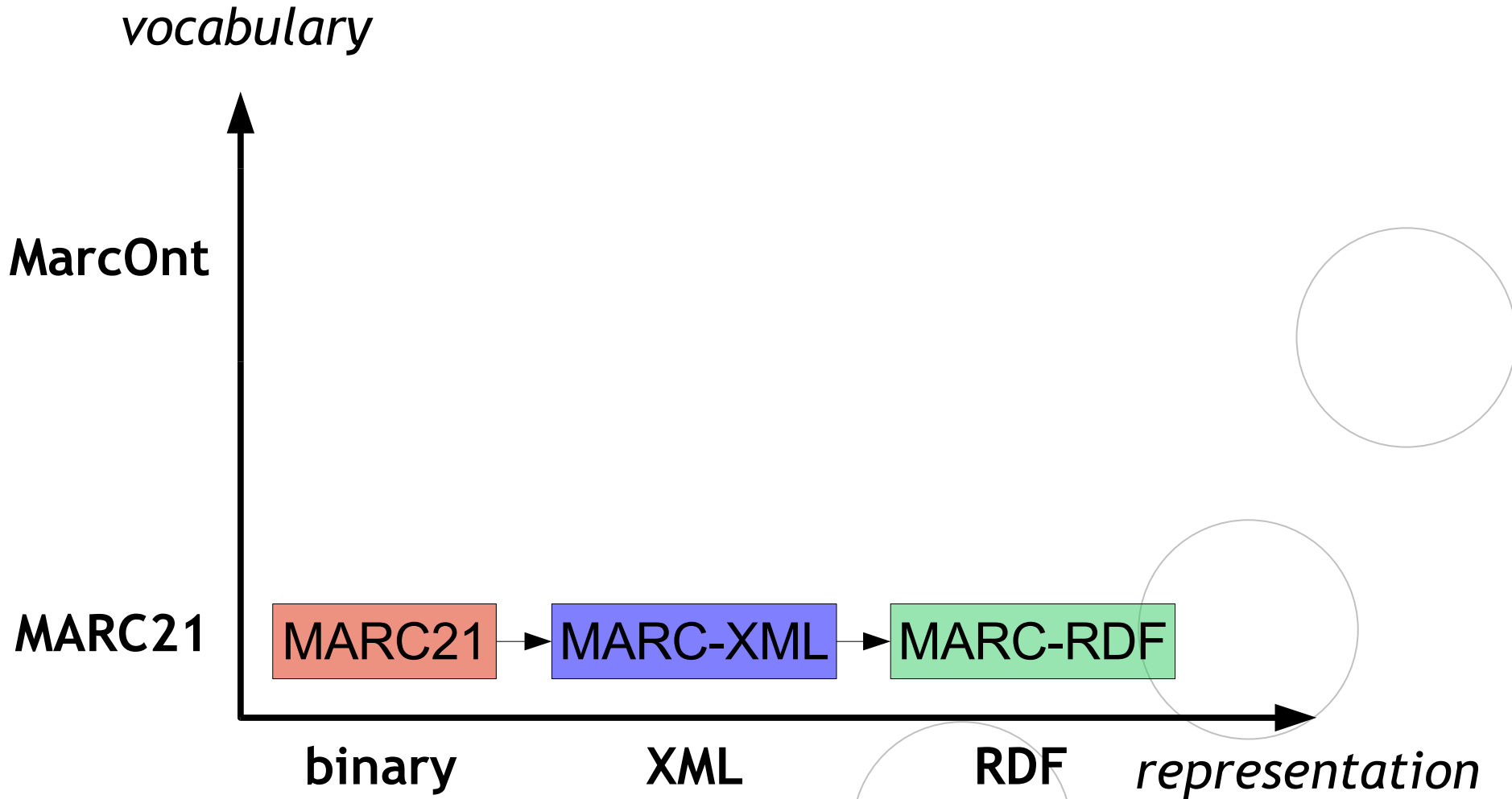




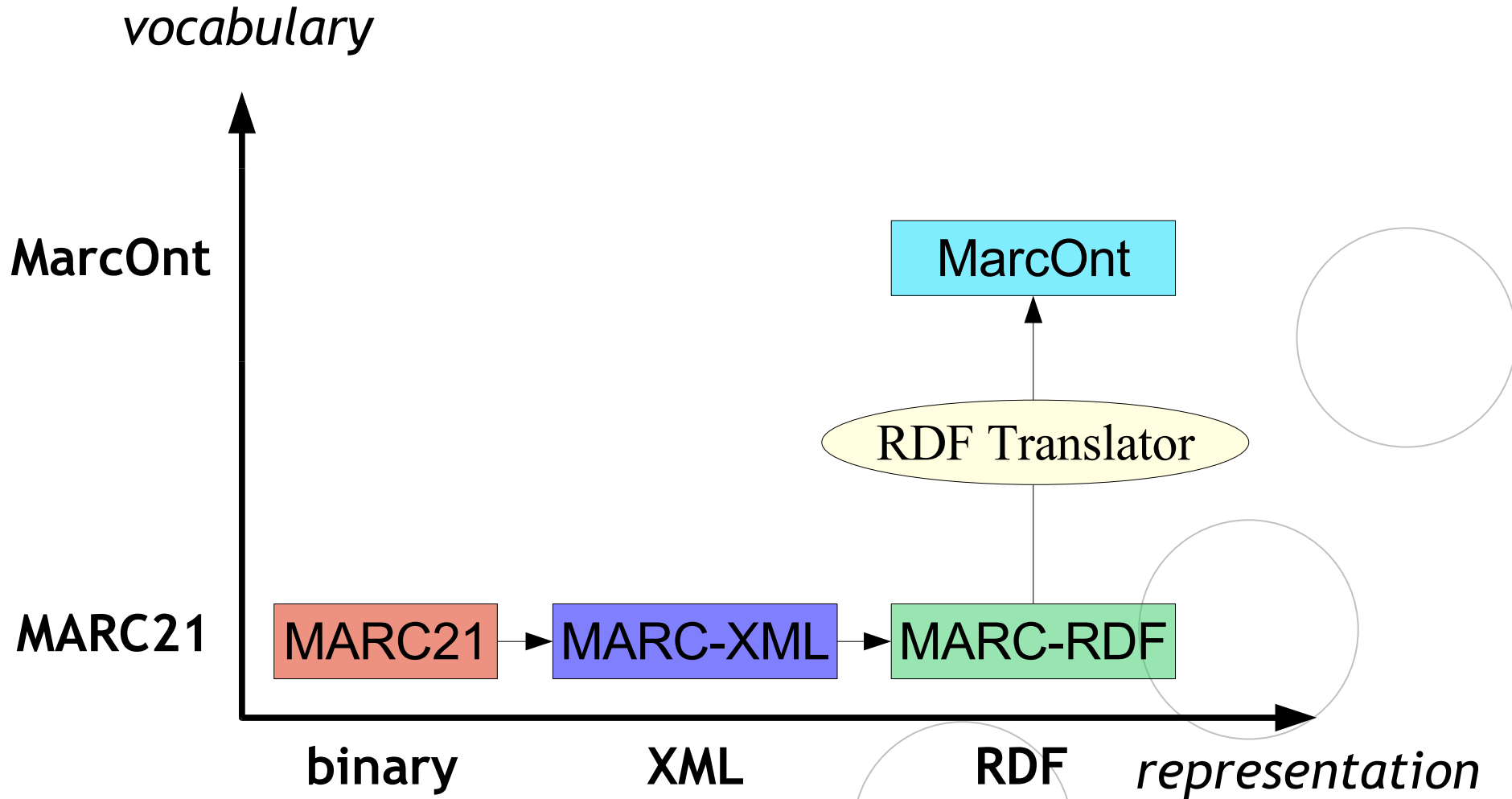
# From Legacy Description to MarcOnt Ontology

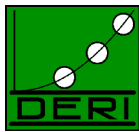


# From Legacy Description to MarcOnt Ontology



# From Legacy Description to MarcOnt Ontology





- All resources are described in **MarcOnt Ontology**, but user can access **MARC21**, **BibTeX** and **DublinCore** descriptions generated on the fly

DERI INTERNATIONAL

### MarcOnt

```
<rdf:RDF>
<rdf:Description rdf:about='mailto:tomas.vitvar@deri.org'>
<rdf:type rdf:resource='http://xmlns.com/foaf/0.1/Person' />
<foaf:surname rdf:datatype='http://www.w3.org/2001/XMLSchema#string'>Tomas Vitvar
</foaf:surname>
</rdf:Description>
<rdf:Description rdf:about='http://www.marcont.org/ontology#8778119'>
<rdf:type rdf:resource='http://www.marcont.org/ontology#TitleStatement' />
<marcont:titleValue rdf:datatype='http://www.w3.org/2001/XMLSchema#string'>D04.01 WSMO, WSMO
and WSMX Working Groups Participation </marcont:titleValue>
</rdf:Description>
<rdf:Description rdf:about='http://library.deri.ie//pages/show.jsp?id=c8af806b'>
<rdf:type rdf:resource='http://www.marcont.org/ontology#Book' />
<marcont:hasCreator rdf:resource='mailto:tomas.vitvar@deri.org' />
<marcont:hasTitles rdf:resource='http://www.marcont.org/ontology#8778119' />
</rdf:Description>
<rdf:Description rdf:nodeID='node10oqnbtcax1'>
<rdf:subject rdf:resource='http://library.deri.ie//pages/show.jsp?id=c8af806b' />
<rdf:predicate rdf:resource='http://www.marcont.org/ontology/marcont.owl#creator' />
<rdf:object rdf:resource='mailto:tomas.vitvar@deri.org' />
<rdf:type rdf:resource='http://www.w3.org/1999/02/22-rdf-syntax-ns#Statement' />
<marcont:order rdf:datatype='http://www.w3.org/2001/XMLSchema#integer'>1 </marcont:order>
</rdf:Description>
</rdf:RDF>
```

BibTeX BibTeXXML BibTeX-RDF Dublin Core **MARC21** MARC XML MARC-RDF MarcOnt

DublinCore

BibTeX

MARC21



- User can select from wide range of description **properties**, defined in **different metadata**, during query building

The screenshot shows a web interface for building queries. At the top, there are two search fields labeled "Search for" with the text "in content" and "in" next to them. Below these fields is a dropdown menu showing a list of properties. The properties are grouped into three sections: DublinCore, BibTeX, and MARC21. The DublinCore section includes "abstract" and "type". The BibTeX section includes "author", "title", "chapter", "series", "year", "address", "number", "month", "editor", "volume", "type", "pages", "booktitle", "note", "edition", and "publisher". The MARC21 section includes "source aud". To the left of the dropdown menu, there are three cyan boxes labeled "DublinCore", "BibTeX", and "MARC21" with arrows pointing to their respective sections in the dropdown. To the right of the dropdown menu, there are checkboxes for "All quer" and "Perform", a "More pa" button, and a "results" label. At the bottom right, there is a "form" button.

Search for  in content

Search for  in

**DublinCore**

- abstract
- type

**BibTeX**

- author
- title
- chapter
- series
- year
- address
- number
- month
- editor
- volume
- type
- pages
- booktitle
- note
- edition
- publisher

**MARC21**

- source aud

results

form



- Performing advanced **RDF** queries

Your query:

```
select * from {a} b {c}
```

Query Language **SeROI** Res

http://www.jeromedl.org/structure#published	"true"^^http://www.w3.org/2001/XMLSchema#boolean
http://www.jeromedl.org/structure#uploadDate	"20050702"^^http://www.w3.org/2001/XMLSchema#integer
http://www.jeromedl.org/structure#abstract	"The core technology area 'work-flow management' or 'process management' process meta-model defines its own set of process definition concepts with process execution semantics. Current practice is to implement a separate interpreter, as many interpreters exist as process meta-models. Various authors have their process-related functionality. As a consequence, the IT standards as well as the required to learn, to use and to maintain several process management procedures. From their viewpoint, it would be tremendously advantageous to develop domain specific products. The state-of-the-art in process management does not provide execution models of all process meta-models."^^http://www.w3.org/2001/XMLSchema#string
http://www.marcont.org/bibtexrdf#hasAuthor	"Simeon Petkov"^^http://www.w3.org/2001/XMLSchema#string
http://www.marcont.org/bibtexrdf#hasTitle	"m3pe light - Multi-Meta-Model Process Execution Environment"^^http://www.w3.org/2001/XMLSchema#string
http://www.marcont.org/bibtexrdf#hasYear	"2005"^^http://www.w3.org/2001/XMLSchema#string
http://www.marcont.org/bibtexrdf#hasSchool	"Information Systems Institute Vienna University of Technology, Digital Enterprise Research Institute Galway"^^http://www.w3.org/2001/XMLSchema#string
http://www.jeromedl.org/structure#href	"http://www.w3.org/2001/XMLSchema#string
http://www.jeromedl.org/structure#usageCount	"27"^^http://www.w3.org/2001/XMLSchema#integer
http://www.jeromedl.org/structure#preprint	"true"^^http://www.w3.org/2001/XMLSchema#boolean
http://www.jeromedl.org/structure#published	"true"^^http://www.w3.org/2001/XMLSchema#boolean
http://www.jeromedl.org/structure#uploadDate	"20050803"^^http://www.w3.org/2001/XMLSchema#integer
http://www.jeromedl.org/structure#abstract	"The contemporary Internet offers various services ranging from electronic mail to on-line services. However, most of the authentication and user registration process must be repeated each time from the beginning, requiring constraints. Very often, user management systems do not allow user to view or actual information gathered about them after registration. To overcome the problem, Microsoft Passport have been proposed. In this article we elaborate on potential access to profile data. We present how required security levels in a user management service. We define how the potential user can benefit from this user management service. We present the D-FOAF, a distributed user management system based on FOAF solutions for secure distributed user management system."^^http://www.w3.org/2001/XMLSchema#string

- Different groups of users require different descriptions of resources
- To provide interoperability in heterogeneous networks we need to be able to translate between different metadata without loss of information
- MarcOnt ontology aim to deliver semantic version of existing description metadata
- MarcOnt Mediation Services can translate between various bibliographic description formats
- MMS has been successfully deployed in JeromeDL

- Build portal for collaborative MarcOnt ontology management and development
- Provide full support for MARC21, BibTeX and DublinCore in MarcOnt ontology and mediation rules
- Deploy MarcOnt ontology in BRICKS project
- Build network of heterogeneous networks based on the MarcOnt ontology



# Thank you for your attention

Sebastian Ryszard Kruk

<http://www.MarcOnt.org/>

