

The Dublin Core *1:1 Principle* in the Age of Linked Data

Dr. Richard J. Urban
Florida State University
School of Information
<http://chi.cci.fsu.edu>
@musebrarian



Acknowledgements

- Urban, R. J. (2012). *Principle Paradigms: Revisiting the Dublin Core 1:1 Principle* (Dissertation). University of Illinois at Urbana-Champaign, Retrieved from <http://hdl.handle.net/2142/31109>
- Dissertation Committee:
 - Dr. Michael B. Twidale
 - Dr. Allen Renear
 - Dr. Carole Palmer
- UIUC Conceptual Foundations Research Group
 - Dr. Karen Wickett
- IMLS Digital Collections and Content Project



1:1

The Principle Defined

In general, Dublin Core metadata describes one manifestation or version of a resource, rather than assuming that manifestations stand in for one another. For instance, a jpeg image of the Mona Lisa has much in common with the original painting, but it is not the same as the painting. As such the digital image should be described as itself, most likely with the creator of the digital image included as a Creator or Contributor, rather than just the painter of the original Mona Lisa. The relationship between the metadata for the original and the reproduction is part of the metadata description, and assists the user in determining whether he or she needs to go to the Louvre for the original, or whether his/her need can be met by a reproduction (Hillmann, 2003).

The Principle Defined

In general, Dublin Core **metadata describes one manifestation or version of a resource**, rather than assuming that manifestations stand in for one another. For instance, **a jpeg image** of the Mona Lisa has much in common with **the original painting**, but it is **not the same** as the painting. As such the digital image should be described as itself, most likely with the creator of the digital image included as a Creator or Contributor, rather than just the painter of the original Mona Lisa. **The relationship** between the metadata for the original and the reproduction **is part of the metadata description**, and assists the user in determining whether he or she needs to go to the Louvre for the original, or whether his/her need can be met by a reproduction (Hillmann, 2003).

“One to one...is a many-headed-snake



and it has bitten us often over the years”

-Stu Waibel

How did we get the *1:1 Principle*?



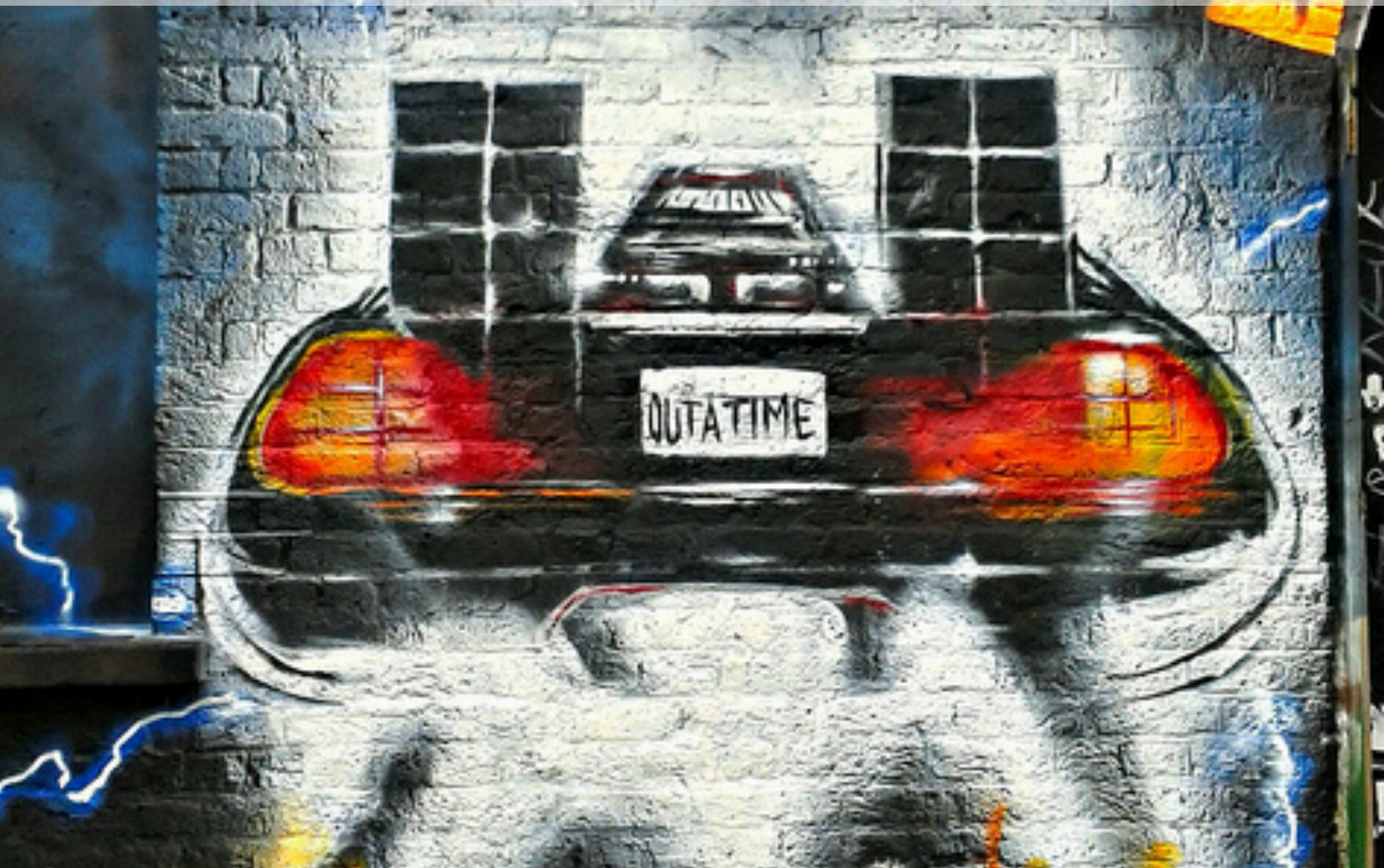
How did we get the *1:1 Principle*?



How did we get the *1:1 Principle*?



How did we get the *1:1 Principle*?



What are we describing? (Simonton, 1962)

Facsimile Theory

- The ***Facsimile Theory*** privileged the intellectual content of an item by making the “original” resource the focus of the record representing a reproduction. Following the long-standing practice of dash entries, a description of the reproduction itself would be included as a note.
- **AACR1**
- **LCRI 1.11A**

Edition Theory

- The ***Edition Theory*** required a record to represent the physical features of the reproduction, using a note to provide a description of the “original” resource.
- **AACR2 “Cardinal Principle”** describe the “item-at-hand”

See Miller (2010) in regards to how these can be applied to 1:1 Principle problems/solutions. <http://dcpapers.dublincore.org/ojs/pubs/article/view/1043/992>

TITLE	Uno mas uno [microform]
PUBLISHED	La Jolla : University of California, San Diego, 1985-
DESCRIPTION	reels : ill. ; 35 mm.

Figure 1. Microform cataloged according to AACR2, Chapter 11

TITLE	Uno mas uno [microform]
PUBLISHED	Mexico, D.F. : Editorial Uno S.A. de C.U., 1978-
DESCRIPTION	v. : ill. ; 46 cm.

Figure 2. Microform cataloged according to LC rule
interpretation

Beyond the Book

Describing Visual Resources and Artworks

- MARC Visual Materials (1980s)
- MARC Archival Manuscript Control (AMC) (1980s)
- ArtMARC (1997)
- Research Libraries Group (RLG)/RLIN (1978)
- Visual Resource Association (1982)
- Getty Art History Information Program (AHIP) (1983-1999)
 - (Later the Getty Information Institute)



Principle Precursor?

- The [*Art and Architecture Thesaurus*] considers **reproductions of works of art to be surrogates for original works** and will recommend that they be indexed in a similar fashion. For example, PAINTING (655) would be used to describe both Leonardo's *Mona Lisa* and a slide reproduction; SLIDE (655) would also be used in the latter case. **This holds serious implications for effective retrieval....In an integrated database containing both of these media, searchers interested only in examples of actual paintings might have to learn to exclude slides, microfilm, and other reproduction media in their search queries to retrieve only records for original paintings.** . . . One solution might be the addition of a “reproduction” facet to indexing strings for object surrogates so that they would be differentiated from “originals” in a browse display

– (Dooley & Zinham, 1990).



Along comes Dublin Core...

- Dublin Core describes “document-like objects” on the Web
 - What counts as a document-like object?
 - What about things that aren’t on the Web that we want to provide access to?
 - What about ***digital reproductions*** and ***surrogates*** of offline ***originals***?

Date	Workshops	Events
1995	DC-1 (Dublin, OH)	
1996	DC-2 (Warwick, GBR)	<ul style="list-style-type: none"> Warwick Framework. How Dublin Core should handle the representation of surrogates left unresolved.
	DC-3 CNI/OCLC Image Workshop (Dublin, OH)	<ul style="list-style-type: none"> Images as DLOs Relationships/Source Discussed
1997	DC-4 (Canberra, AUS)	<ul style="list-style-type: none"> Eric Jul introduces idea for 1:1 “to each resource, it’s own description”
	RLG/ALCTS Summit (Mountain View, CA)	<ul style="list-style-type: none"> Drafting <i>Guidelines for Extending Dublin Core</i>
	DC-5 (Helsinki, FIN)	<ul style="list-style-type: none"> <i>1:1</i> and <i>Relation</i> Work groups established AHDS Report
1998	DC-6 (Washington, D.C.)	<ul style="list-style-type: none"> Discussion continues on listservs RDF emerges from W3C Beginnings of DC Abstract Model
1999	DC-7 (Frankfurt, DEU)	<ul style="list-style-type: none"> Bearman et al. (1999) discusses “logical clusters of metadata” used for the INDECS project.
2000	DC-8 (Ottawa, CAN)	<ul style="list-style-type: none"> Relation and One-to-One Working Groups disbanded
2001	DC-9 (Tokyo, JPN)	<ul style="list-style-type: none"> OAI-PMH 1.1 requires simple Dublin Core.
2003	DC-2003 (Seattle, WA)	<ul style="list-style-type: none"> <i>1:1</i> appears in <i>Using Dublin Core</i>

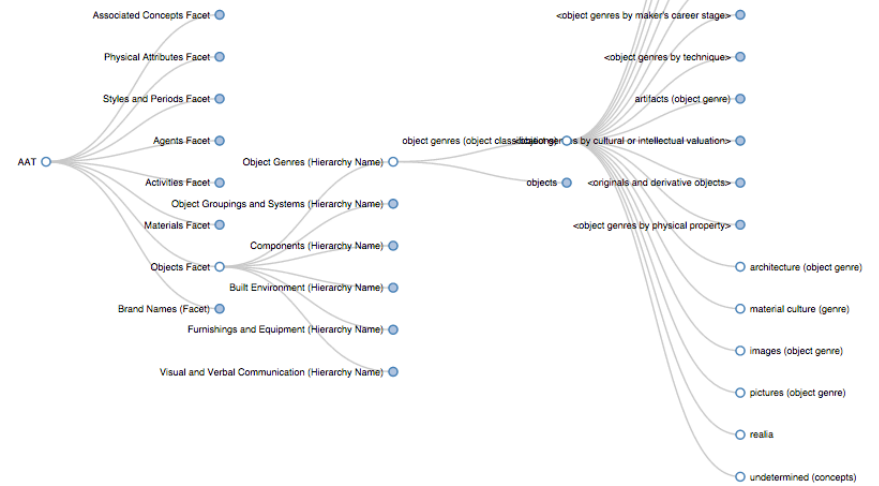
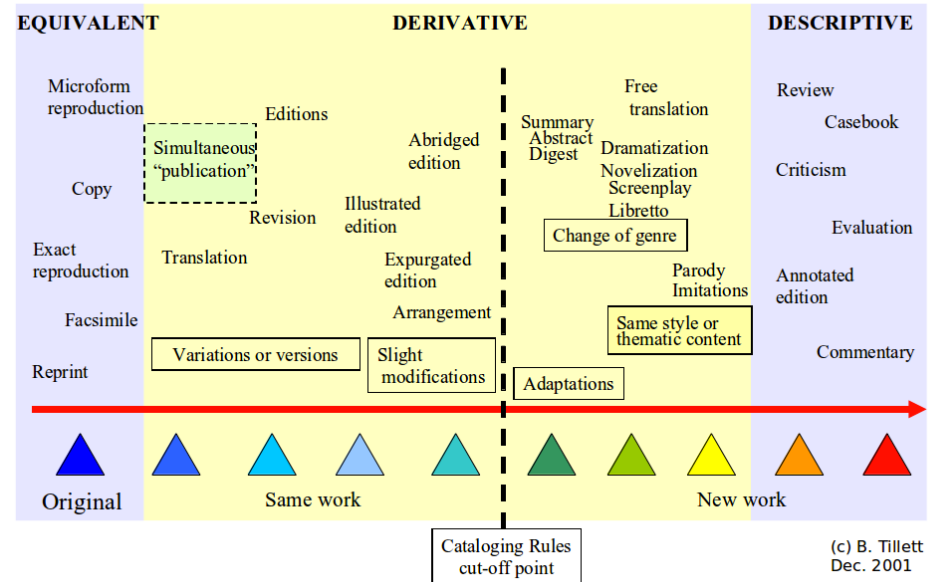
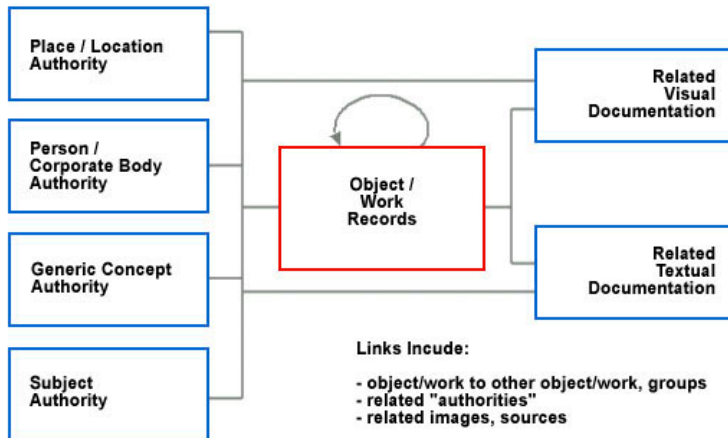
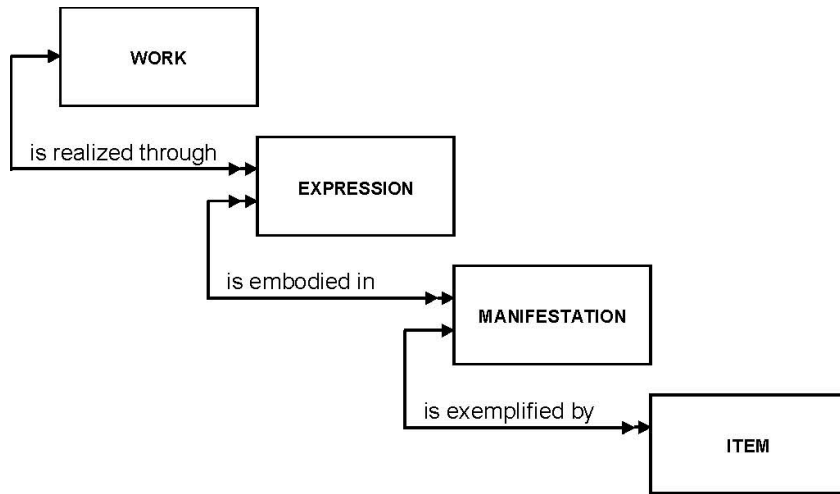
Guidelines for Extending the Use of Dublin Core Elements (RLG, 1997)

- Extend Dublin Core beyond Web-based resources
 - Images as DLOs
 - Non-web resources
- Include “record type indicators”
 - is a DLO the original or not
 - Is the DLO Internet-accessible or not

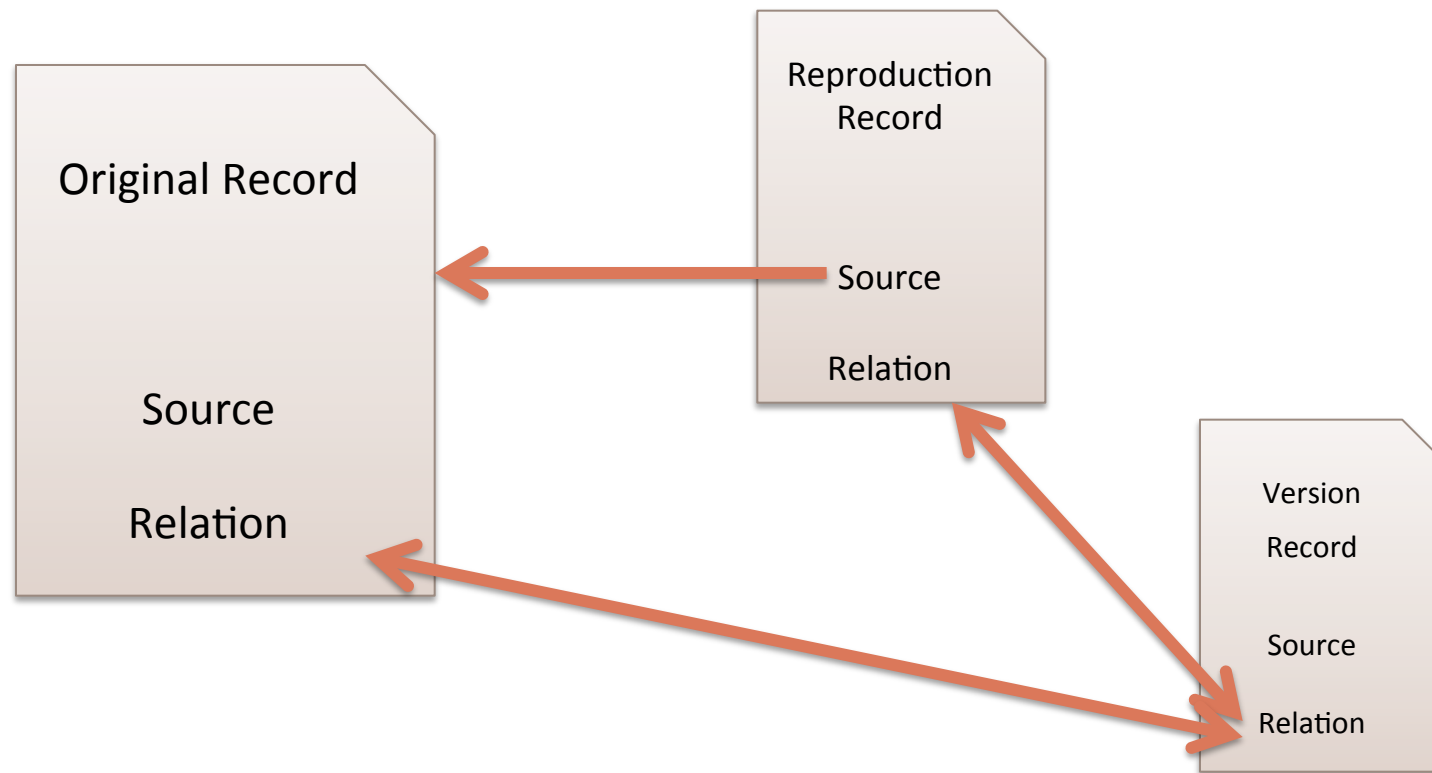
Guidelines for Extending the Use of Dublin Core Elements (RLG, 1997)

- Modify DC term definitions to indicate “values that *describe the intellectual content of the **original resource**.*”
- “Concoct a way to start and end sets of DC elements” in order to describe originals/surrogates.
- Add terms to Resource Type vocabulary for surrogates/offline resources.
- Make better use of Relation/Source to help structure original/surrogate descriptions.

one2one WG



Relations/Source WG

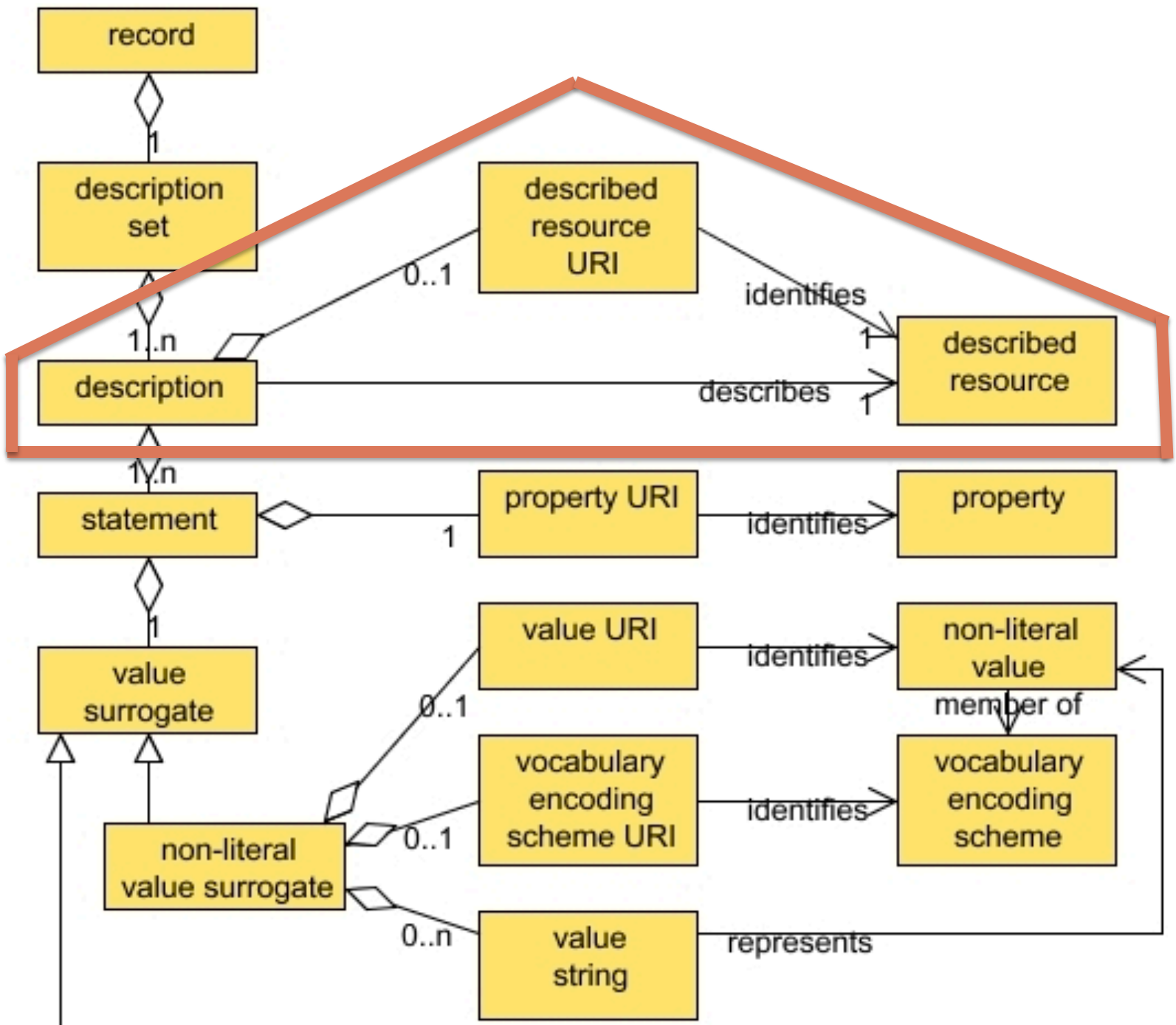


The Principle in DCAM

The abstract model presented above indicates that each DC **metadata description describes one, and only one, resource**. This is commonly referred to as the one-to-one principle.

However, **real-world metadata** applications tend to be **based on loosely grouped** sets of descriptions (where the described resources are typically related in some way), known here as **description sets**. For example, a description set might comprise descriptions of both a painting and the artist. Furthermore, it is often the case that a description set will also contain a description about the description set itself (sometimes referred to as 'admin metadata' or 'meta-metadata').

Description sets are instantiated, for the purposes of exchange between software applications, **in the form of metadata records**, according to one of the DCMI encoding guidelines (for example, XHTML meta tags, XML and RDF/XML) [DCMI-ENCODINGS].



1:1 Principle Violations

<title>Mona Lisa</title>

<date>2008</date>

<date>1501-1519</date>

<source>TIFF</source>

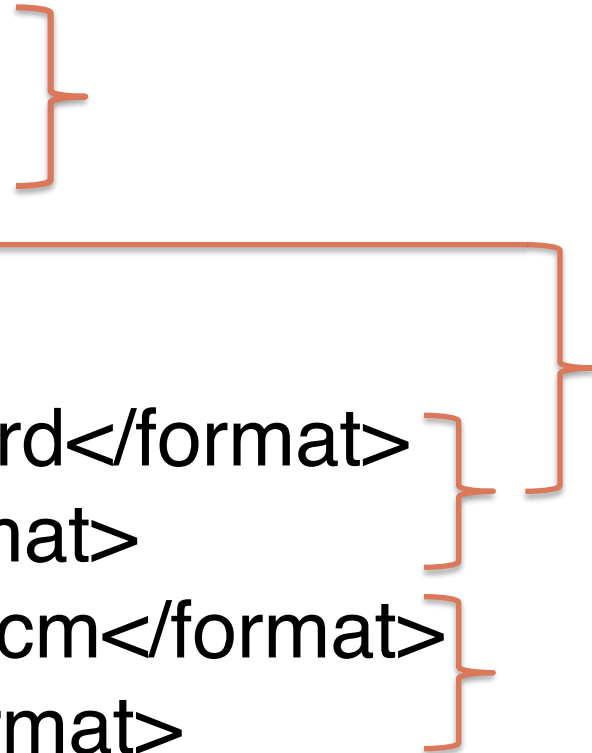
<type>image</type>

<format>oil on poplar board</format>

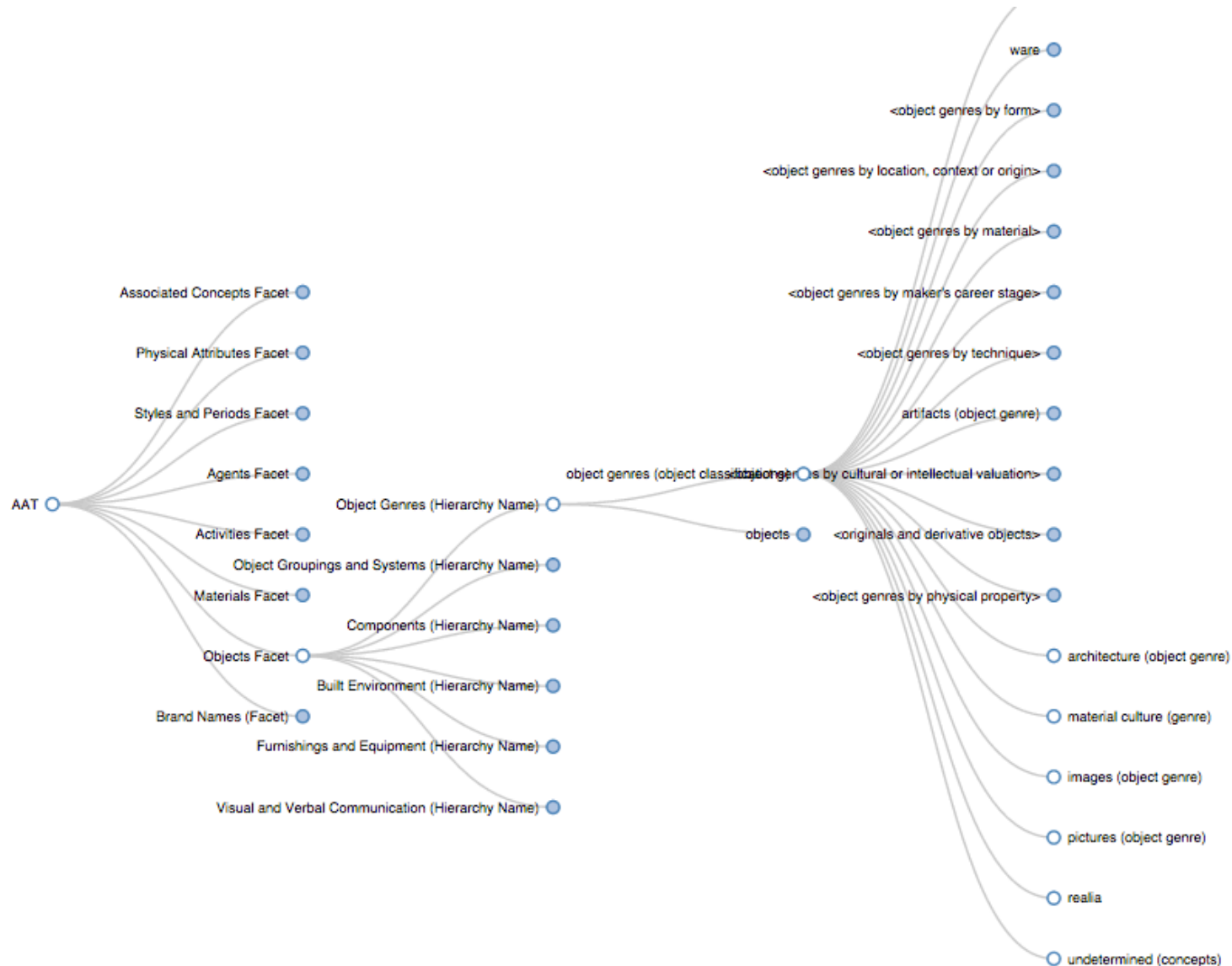
<format>image/jpeg</format>

<format>H. 77 cm; W. 53 cm</format>

<format>16781 bytes</format>



Limitations of Violations



- *resource* - A resource is the object or "stuff" that metadata is "about". The nature of a resource, whether it is physical or digital, or whether it is stored in the repository or is a constituent of another database, is outside the scope of the OAI-PMH.



← resource

item =
identifier

all available metadata
about *David*

← item

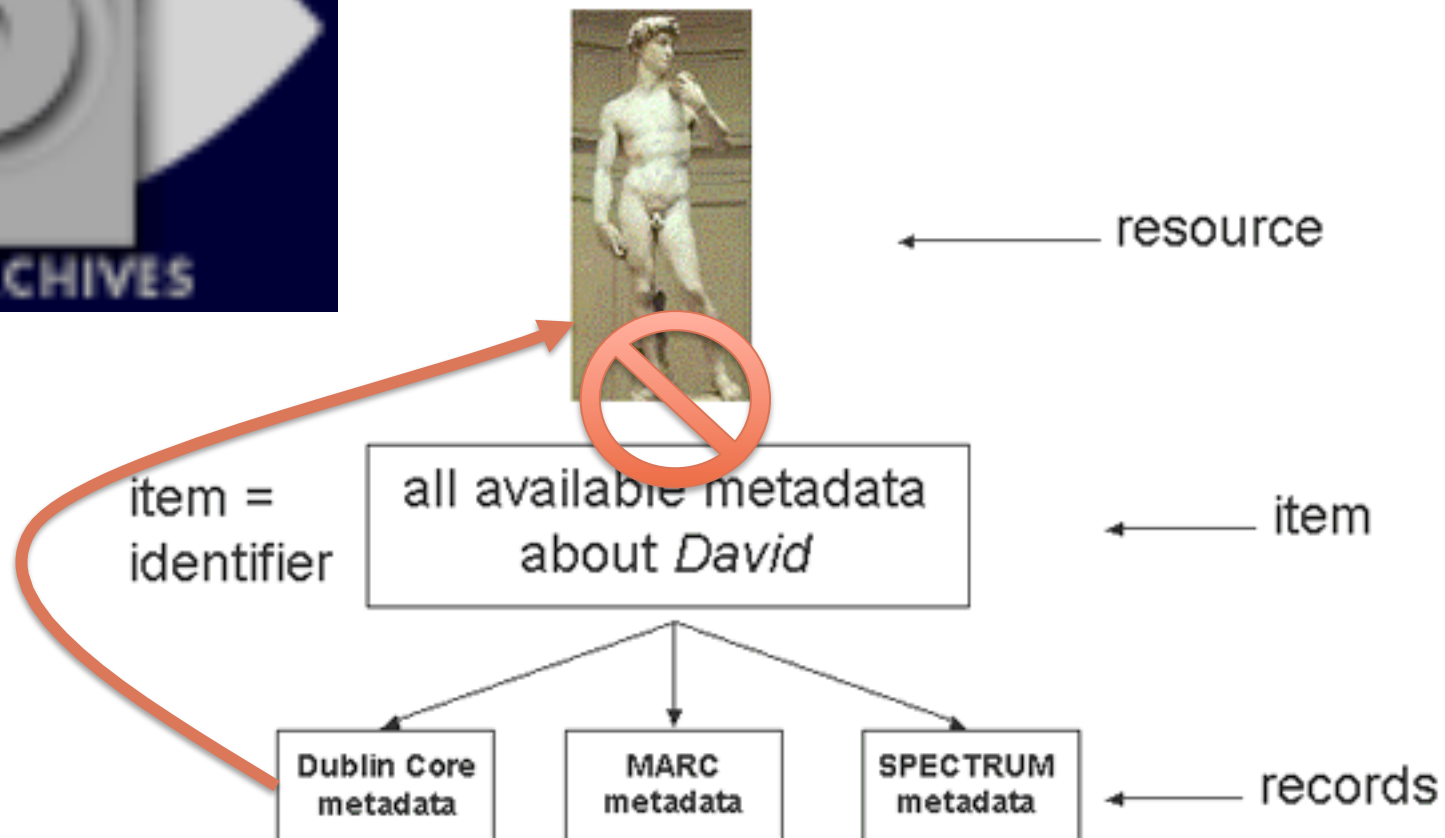
Dublin Core
metadata

MARC
metadata

SPECTRUM
metadata

← records

- *resource* - A resource is the object or "stuff" that metadata is "about". The nature of a resource, whether it is physical or digital, or whether it is stored in the repository or is a constituent of another database, is outside the scope of the OAI-PMH.

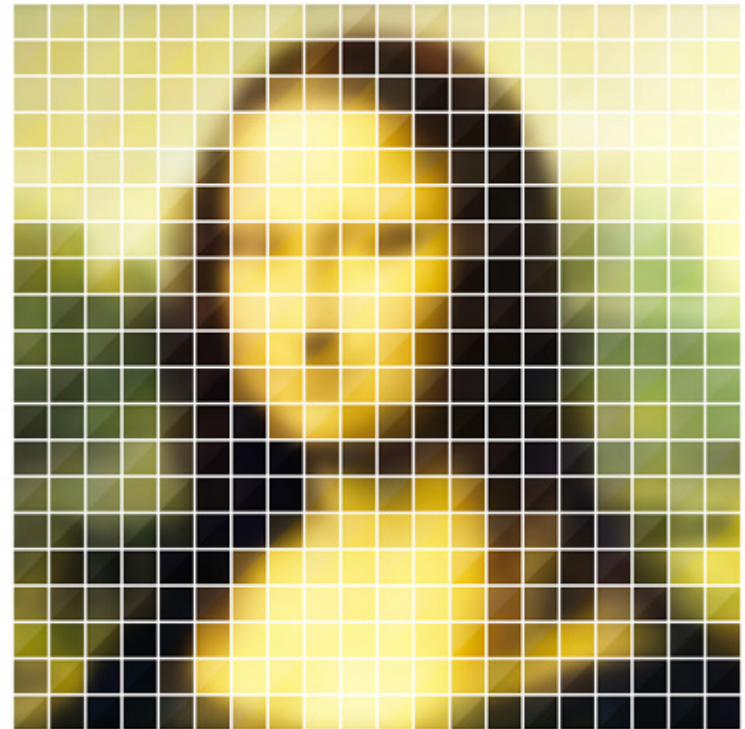


Reference an important part of semantics

date.original
refersTo

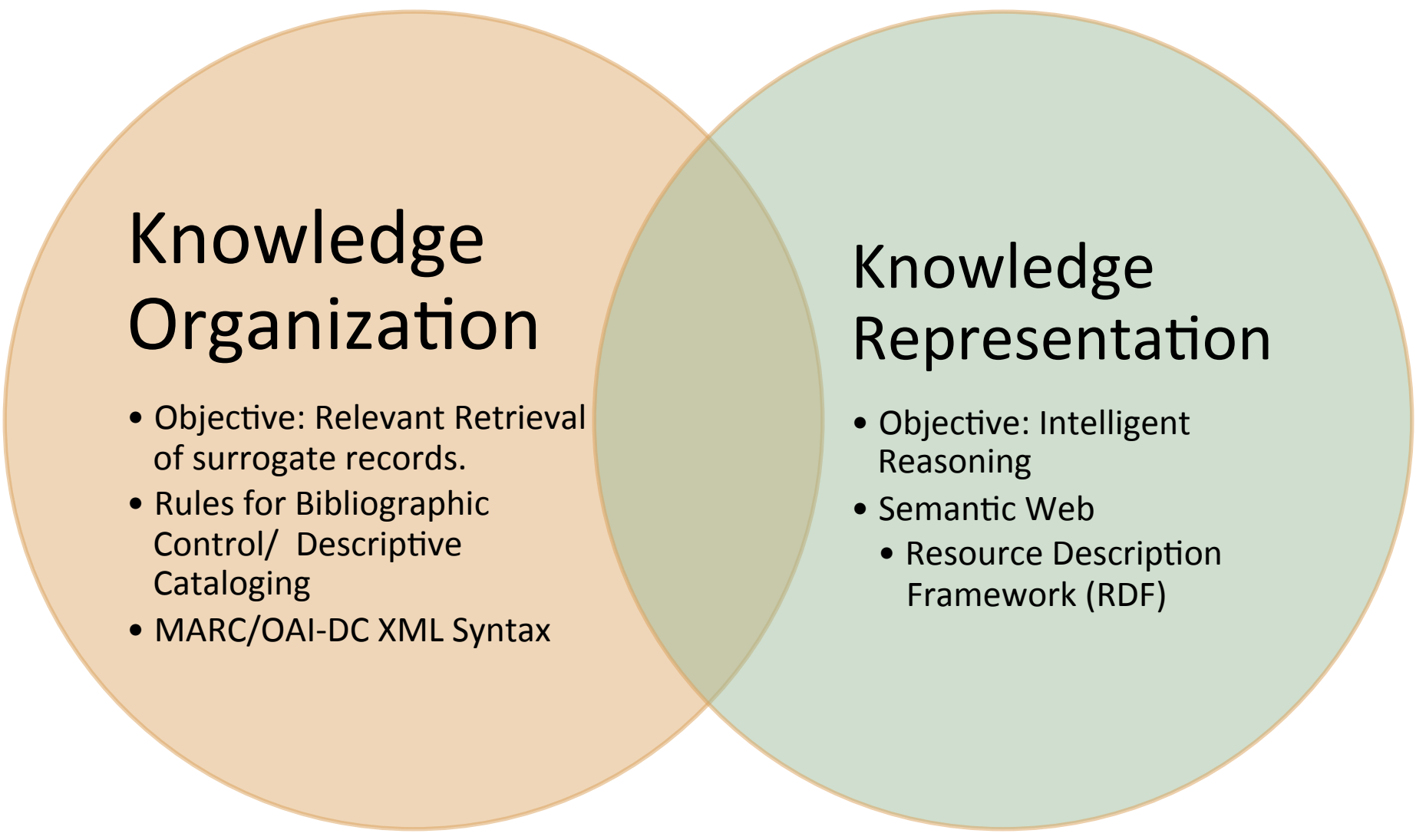


date.digital
refersTo





SAVE US, RDF!



Knowledge Organization

- Objective: Relevant Retrieval of surrogate records.
- Rules for Bibliographic Control/ Descriptive Cataloging
- MARC/OAI-DC XML Syntax

Knowledge Representation

- Objective: Intelligent Reasoning
- Semantic Web
 - Resource Description Framework (RDF)

Knowledge Organization

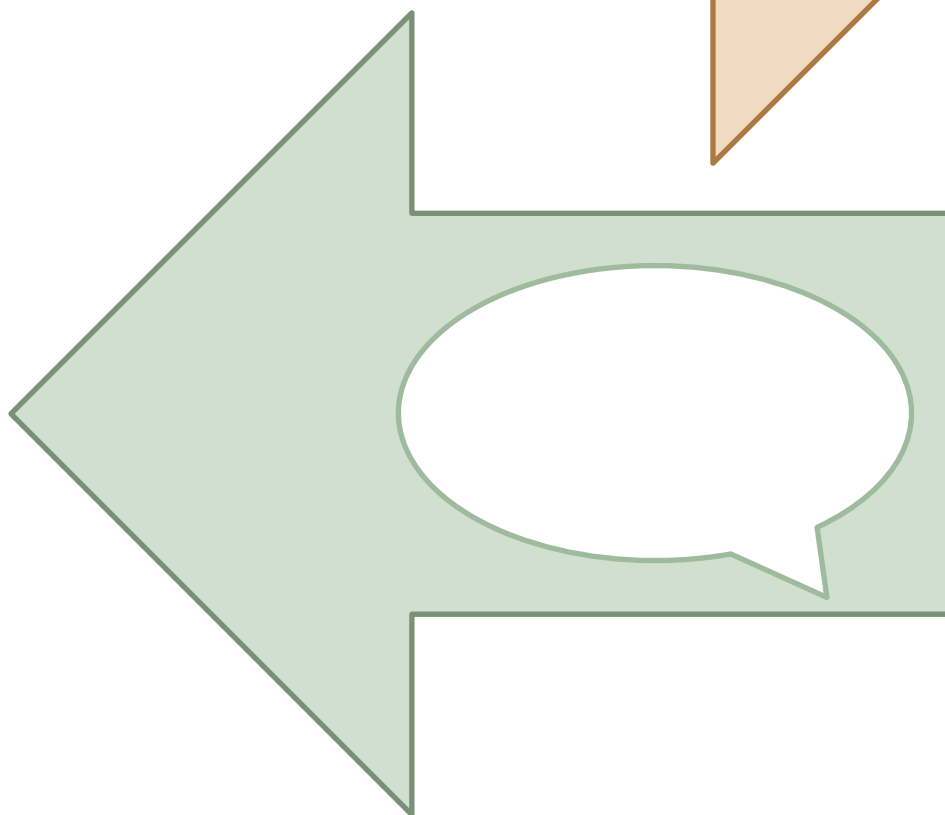
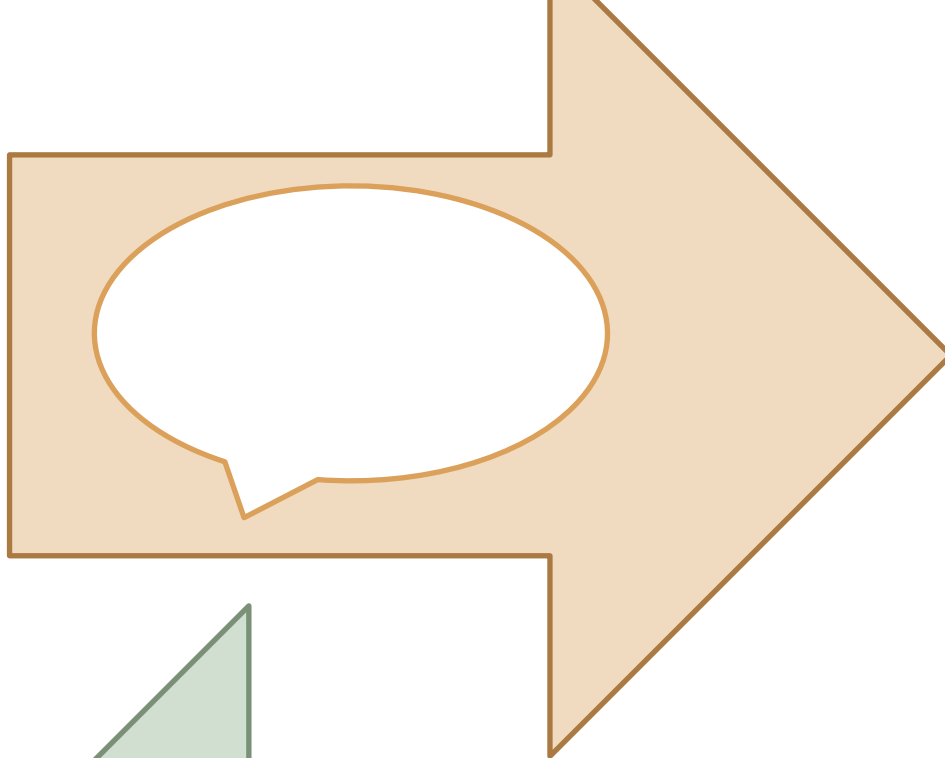
- Identity through definite descriptions
- Fixed record syntaxes with informal semantics.
- *Meaning* supplied by principles, rules, etc.

1:1

Knowledge Representation

- Identity through names (URIs)
- Complex graph of relationships among named resources.
- Formal semantics supplied by *interpretations* and *models*.

KO



KR

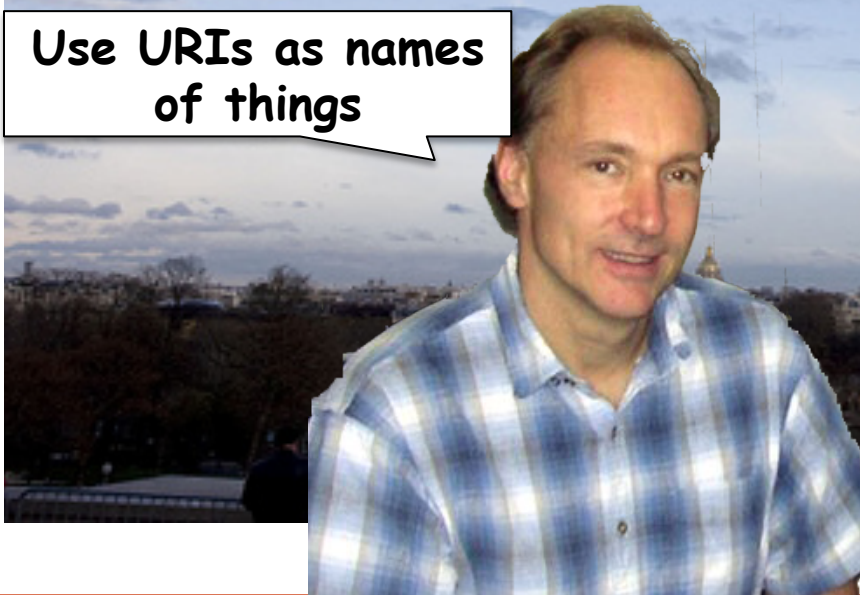
`http://ex.org/eiffelTower`

refers to

creator: Gustav Eiffel

date: 1889

**Use URIs as names
of things**



<http://ex.org/eiffelTower>

refers to

creator: Gustav Eiffel

date: 1889

Use URIs as names
of things



http://ex.org/eiffelTower

Content
negotiation

<http://dbpedia.org/resource/Eiffel_Tower> <http://www.w3.org/2002/07/owl#sameAs> <http://rdf.freebase.com/ns/m.02j81> .
<http://dbpedia.org/resource/Eiffel_Tower> <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://dbpedia.org/class/yago/Skyscraper104233124> .
<http://dbpedia.org/resource/Eiffel_Tower> <http://dbpedia.org/property/id> "s0000021"@en .
<http://dbpedia.org/resource/Eiffel_Tower> <http://www.w3.org/2002/07/owl#sameAs> <http://cy.dbpedia.org/resource/T\u0175r_Eiffel> .
<http://dbpedia.org/resource/Eiffel_Tower> <http://dbpedia.org/property/caption> "The Eiffel Tower as seen from"@en .
<http://dbpedia.org/resource/Eiffel_Tower> <http://www.w3.org/2002/07/owl#sameAs> <http://ilo.dbpedia.org/resource/Torre_Eiffel> .
<http://dbpedia.org/resource/Eiffel_Tower> <http://www.w3.org/2002/07/owl#sameAs> <http://io.dbpedia.org/resource/Eiffel_Turmo> .

**Use URIs as names
of things**



owl:sameAs?

```
<http://dbpedia.org/resource/Eiffel_Tower>  
  <http://www.w3.org/2002/07/owl#sameAs>  
  <http://rdf.freebase.com/ns/m.02j81> .  
<http://dbpedia.org/resource/Eiffel_Tower>  
  <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>  <http://dbpedia.org/class/  
yago/Skyscraper104233124> .  
<http://dbpedia.org/resource/Eiffel_Tower>  
  <http://dbpedia.org/property/id>  
  "s0000021"@en .
```



What does this mean for RDF Validation?

UC-11

alphanumeric ID: UC-10-WRONG-PROPERTY-DOMAINS

label: Wrong Property Domains

description:

Check if the (datatype or object) property is in the right domain.

case studies:

CS-3

CS-6

requirements:

R-25

R-26

What does this mean for RDF Validation?

UC-Europeana-15

alphanumeric ID: UC-Europeana-15-URI-target-validation

label: URI target validation

description:

Some EDM properties should be used in combination with URI pointing to specific media. Some constraints might be defined for these medias (file type, size) .

edm:isShownBy: Check that the URI resolves to a media

edm:object: Check that pointed media respects the requirements

<http://pro.europeana.eu/documents/900548/960640/Europeana+Portal+Image+P...>

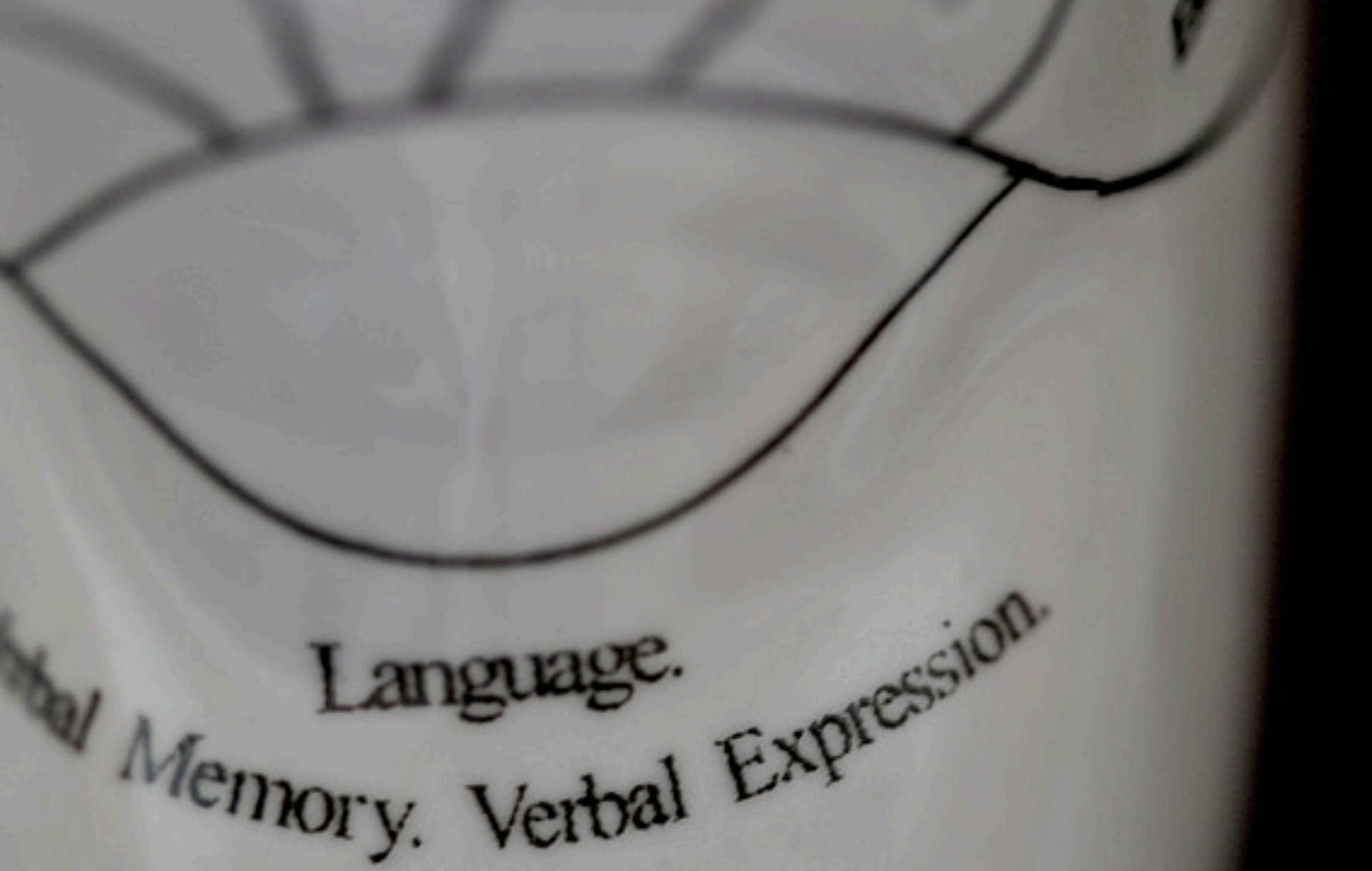
case studies:

[CS-3](#)

[CS-4](#)

requirements:

[R-171](#)



How can we understand and define
metadata mindsets.

Credits

- Portrait of Lisa Gherardini, wife of Francesco del Giocondo, known as the Mona Lisa (the Joconde in French) © Musee du Louvre/A. Dequier – M. Bard
- Double Headed Snake - British Museum by <http://www.flickr.com/photos/dunechaser/2793043767/>
- The DeLorean Time Machine in "Back to the Future"
<https://commons.wikimedia.org/wiki/File:Delorean5.jpg>
- Becraft, Andrew. Lego Back to the Future Delorean <https://flic.kr/p/fbD2xD>
- Gonzalez, Jorge Philipe. Maquinas del Tiempo. <https://flic.kr/p/4hxhib>
- MsSarahKelly. Back to the Future by Graffiti Life. <https://flic.kr/p/fmBsBt>
- flying_rhino_303 Andy Warhol "30 Mona Lisa's" Pop Art Conceptual Art artist 35mm slide
- http://www.ebay.com/itm/Andy-Warhol-034-30-Mona-Lisa-039-s-034-Pop-Art-Conceptual-Art-artist-35mm-slide-/291197617937?nma=true&si=dWZ0WmtfQUo9bQaQ%252BB7jDrq4slo%253D&orig_cvip=true&rt=nc&trksid=p2047675.l2557
- Tillett, B. Horizontal Bibliographic Relationships
https://pantherfile.uwm.edu/kipp/public/courses/511/511notes-bibstructs_html_m42e2c6da.png
- CDWA E-R Diagram http://www.getty.edu/research/publications/electronic_publications/cdwa/entity.html
- Lincoln, Matthew Hierarchies of AAT <http://matthewlincoln.net/pages/dendrogram-fullscreen.html>
- Pixel Mona Lisa <http://www.whatkeptmefrom.com/blog/?q=pixelartseries>
- Stars Wars Hologram <http://static.thetechjournal.net/wp-content/uploads/2012/10/star-wars-hologra.jpg>
- Stopethegears <https://flic.kr/p/2F4zSb>