Modelling CDWA Lite as an OWL-DL Ontology

Richard J. Urban

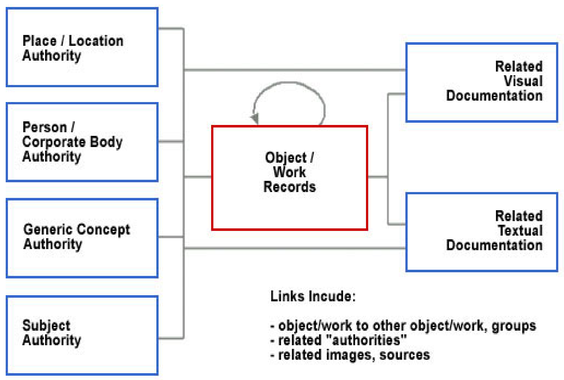
rjurban@illinois.edu http://isrl.uiuc.edu/~rjurban

GRADUATE SCHOOL OF LIBRARY AND INFORMATION SCIENCE
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
The iSchool at ILLINOIS

Introduction

The Categories for the Description of Works of Art "describes the content of art databases by articulating a conceptual framework for describing and accessing information about works of art, architecture, other material culture groups and collections of works and related images." Developed by the Art Information Task Force (AITF) in the early 1990s, CDWA has served as a foundational framework for the description of cultural heritage materials. In order to facilitate sharing of CDWA compliant records via the Open Archives Initiative - Protocol for Metadata Harvesting (OAI-PMH) the CDWA Lite XML schema, based on CDWA "core" categories was created in 2006. [1]

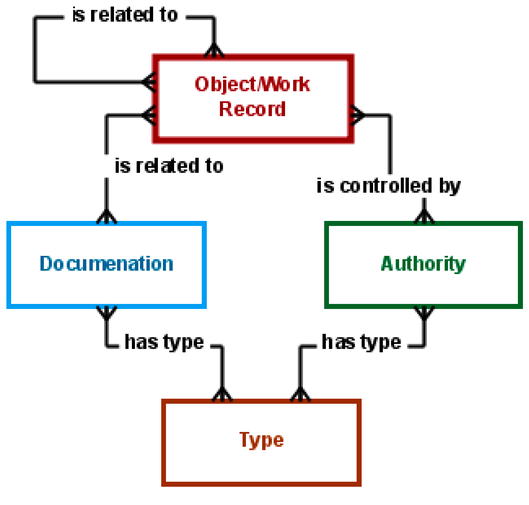
The CDWA "data model"



http://www.getty.edu/research/conducting_research/standards/cdwa/entity.html

Above is the "E-R Diagram" that frequently appears in literature about the CDWA. While this diagram represents several important entities found in the CDWA, it lacks other features of functional E-R diagrams, such as clearly named relationships between entities, or indications of cardinality.

Below is an improved E-R diagram that posits general "documentation" and "authority" entities that are modified by a "type" entity. While it fixes some of the problems displayed in the diagram above, it is still not a offer a full representation of the entities and relationships found in the CDWA Lite XML Schema.

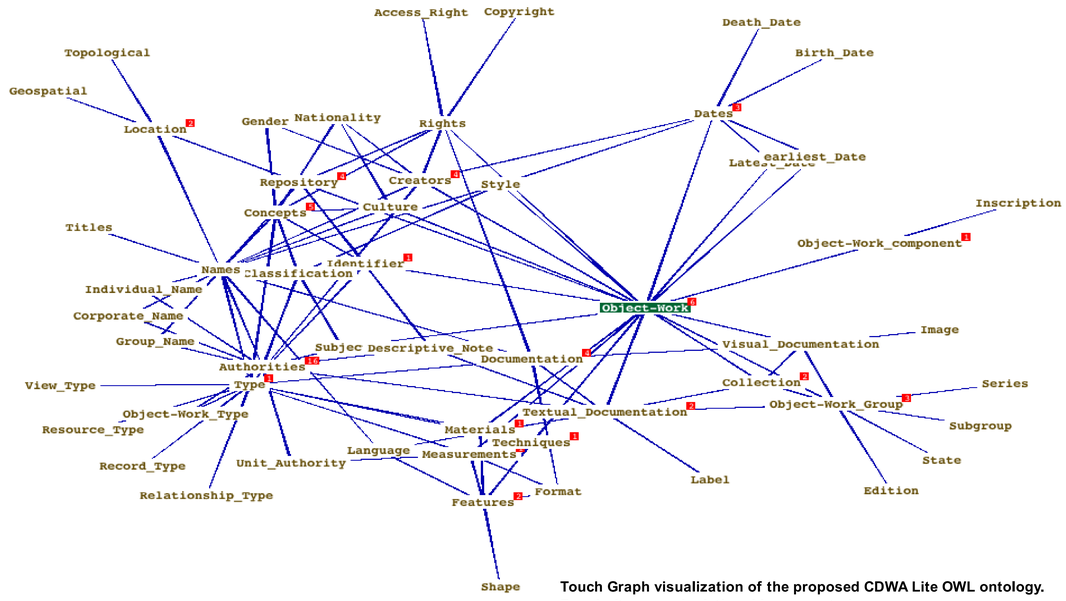


Building a CDWA Lite Ontology

In order to make CDWA compliant data available on the emerging Semantic Web, it will be necessary to go beyond the syntax of the CDWA Lite XML schema towards representations in the Web Ontology Language (OWL). For this project CDWA Lite was modelled using Standford's Protégé ontology editor and Noy & McGuiness' "Ontology 101" method. [3] [4]

Although CDWA tries to avoid questions about the ontology of art, it does contain some implicit conceptualizations of the domain which presents interesting challenges when building an OWL model.

For example, CDWA Lite records may also describe *collections*, *series* and *groups of works* which are not represented in the broader model. Does this mean they are sub-classes of Object-Works or a disjoint class? Similarly, CDWA/CDWA Lite make a distinction between entities that are *Object/Works* and things that are *Related Resources* or *Visual/Textual Documentation*. While it is possible to create a generalized class of features (such as measurements, formats, titles, locations, etc.) that both kinds of things exhibit, CDWA restricts them to only being features of *Object/Works*.



Touch Graph visualization of the proposed CDWA Lite OWL ontology. For clarity, named relationships are not shown

Why OWL-DL?

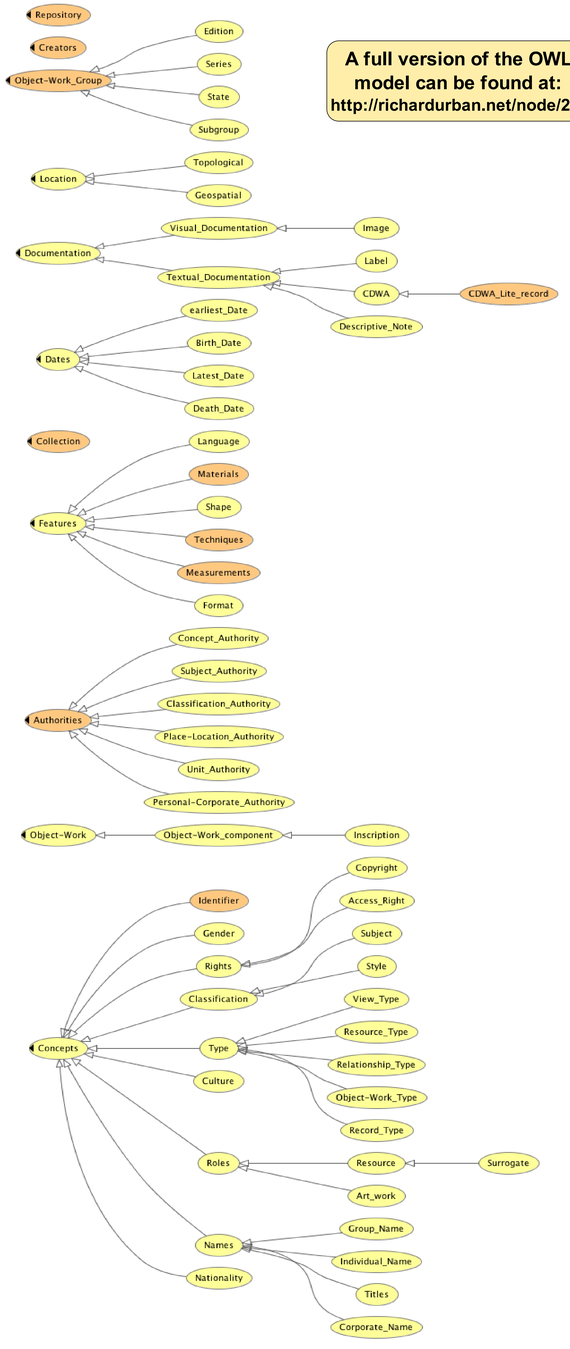
Becasue CDWA Lite is interested in talking about *Object/Works* and *Resources/Documentation* as disjoint classes of things that each have distinct properties, the expressive power of OWL-DL is needed for a CDWA Lite Ontology. While it may be possible to express some parts of CDWA Lite using just RDF Schema (RDFS), it would be necessary for the CDWA community to re-imagine its approach to modeling both *works of art* and *resrouces*. One possible solution is to add a class of *Roles* that distinguishes between descriptions of *works of art* and its related documentation.

Future Directions

Currently the CIDOC Conceptual Reference Model (CIDOC-CRM) provides the richest domain ontology for describing cultural heritage materials. As a model for data interchange and harmonization, the CRM has been used to map many of the metadata standards used by the cultural heritage sector, including MARC, Dublin Core, EAD. Because CIDOC-CRM can be daunting in its complexity and might have contained assumptions about works of art that are not shared by CDWA, this exercise choose to allow CDWA Lite to speak for itself. However, some of the problems noted above suggest that a CDWA-CRM mapping could provide useful suggestions for improving CDWA. The MuseumDAT project has already demonstrated that CIDOC-CRM can be a useful tool for improving the CDWA Lite XML schema. [2] A CDWA ontology model informed by CIDOC-CRM would strengthen CDWA by refining and clarifying muddled class and property concepts. This may, however, also require the communnity to rethink how definitions of CDWA are written and how to resolve places where CDWA concepts do not fit in the CIDOC-CRM.

CDWA has also influenced related standards such as the Visual Resource Association *VRACore* and the *Cataloging Cultural Objects* content standard. Treating these as a suite of complimentary resources, rather than relying on any one representation, may result in an interoperable model that will be useful across both communities.

CDWA Lite OWL classes



A full version of the OWL model can be found at: http://richardurban.net/node/278

References

[1] J. Paul Getty Trust and College Art Association. 2006. *Categories for the Description of Works of Art.* http://www.getty.edu/research/conducting_research/standards/cdwa/index.html

[2] Stein, R. and Coburn, E. 2008. CDWA and MuseumDAT: New Developments in Metadata Standards for Cultural Heritage. The annual conference of the International Documentation Committee of the International Council of Museums. Athens, 15-18 September 2008. http://www.cidoc2008.gr/cidoc/Documents/papers/drfile.2008-06-17.0283568160

[3] Noy, N.F. and McGuinness, D.L. 2001. Ontology Development 101: A Guide to Creating Your First Ontology http://ksl.stanford.edu/people/dlm/papers/ontology101/ontology101-noy-mcguinness.html

[4] Horridge, M., Knublauch, H, Rector, A., Stevens, R., and Wroe, C. 2004. A Practical Guide to Building OWL Ontologies Using the Protégé-OWL Plugin and CO-ODE Tools. http://www.co-ode.org/resources/ tutorials/ProtegeOWLTutorial.pdf