# Towards an Ontology for Art and Colours

## Luciana Bordoni, Tiziana Mazzoli

ENEA – UDA/Advisor

Via Anguillarese 301, 00060 S. Maria di Galeria (Rome), Italy

bordoni@casaccia.enea.it, tiziana_mazzoli@yahoo.it

### Abstract

To meet a variety of needs in information modeling, software development and integration as well as knowledge management and reuse, various groups within industry, academia, and government have been developing and deploying sharable and reusable models known as ontologies. Ontologies play an important role in knowledge representation. In this paper, we address the problem of capturing knowledge needed for indexing and retrieving art resources. We describe a case study in which we attempt to construct an ontology for a subset of art. The aim of the present ontology is to build an extensible repository of knowledge and information about artists, their works and materials used in artistic creations. Influenced by the recent interest in colours and colouring materials, mainly shared by French researchers and linguists, an ontology prototype has been developed using Protégé. It allows to organize and catalogue information about artists, art works, colouring materials and related colours.

## 1. Introduction

An important requirement for the semantic analysis of textual documents is the annotation of terms with semantic classes in areas such as document classification, information retrieval and information extraction. This allows a semantic normalization of different terms to the same semantic class (concept). Semantic lexicons, like synonym lists (Roget), and lexical semantic databases derived from machine-readable dictionaries (LDOCE, Webster, among others) or built from scratch (WordNet, EuroWordNet, and similar resources) provide information on semantic classes. However, a number of problems hampers the employment of such resources in real world environment (mostly domain specific). Meanings are often represented in the resource that are of no relevance to the specific domain under consideration. Whereas, on the other hand, domain specific terms often carry meanings that are not represented in the resource, because either the meaning is not associated with this term or the meaning is not available at all.

To meet a variety of needs in information modeling, software development and integration as well as knowledge management and reuse, various groups within industry, academia, and government have been developing and deploying sharable and reusable models known as ontologies.

Ontologies represent many different kinds of things in a given subject area. These things are represented in the ontology as classes (sometimes called concepts) and are typically arranged in a lattice or taxonomy of classes and subclasses. Each class is usually associated with various properties describing its features and attributes as well as various restrictions on them.

The promise of ontologies is "*a shared and common understanding of a domain that can be communicated between people and application systems*" (Fensel, 2001). Recently, more and more projects and theoreticians support the use of formal ontologies as common schema for information integration. In this paper we address the problem of capturing knowledge needed for indexing and retrieving art resources. We describe a case study in which we attempt to construct an ontology for a subset of art. Influenced by the recent interest in colours and colouring materials, mainly shared by French researchers and linguists, an ontology prototype has been developed to organize and catalogue information about artists, art works, colouring materials and related colours. The aim of the present ontology is to build an extensible repository of knowledge and information about artists, their works and materials used in artistic creations. The remainder of the article is organized as follows. Section 2 presents the ontology design process. Section 3 explains about the ontology and its rule building. Section 4 discusses conclusion and possible future works.

## 2. Ontology design process

The research work has started by consulting existing thesauri such as the Art and Architecture Thesaurus[1] (AAT). It presents a single inheritance hierarchy of categories (Wielinga, 2001) and provides a classification scheme for art and architecture concepts. The purpose of the AAT is not the classification of real art object such as paintings or their authors and no category is instantiated. For this reason, our research pointed to build a category structure apt to include existing art works, artists, related artistic movements and colouring materials which can be classified by related colour hues. A thesaurus on painting and sculpture has been edited to dispose of a reusable database. The planning of the thesaurus itself introduced some of the formal and structural issues the subsequent ontology was about to raise.

One of the concerns in contemporary art classification is the criterion an indexer should follow to assign the right value of genuine work of art to experiences and explorations that have not gained institutional recognition yet. In general, the indexing choice of art instances could profitably comply with the following

---

[1] http://www.getty.edu/research/conducting_research/vocabularies/aat

ISO 21127 description of the Iconographic Object entity: *"This entity comprises objects which are designed primarily or in addition to another functionality to represent or depict something in an optical manner, be it concrete or abstract. This entity has certain pragmatic value in the fine arts since it conveniently groups together objects such as paintings, drawings, watercolours, and other similar objects. From a philosophical point of view, representation is an 'intentional' act. Natural objects may resemble other objects by chance but they can represent only as a result of intervention by some fairly sophisticated semiotic arrangements"*[2]. The standard gives directions about the separation of the artefact from the natural object. It underlies the intentional nature of the human act placing the artefact into the fine arts domain and the resulting classification of the artefact (whose ISO 21127 super class is the Man-Made Object entity) as an Iconographic Object.

Another indexing issue has been the managing of homonymous thesaurus entries belonging to different domains. In fact, even in the art field, where the uniqueness should be a distinctive feature, we can often find this kind of problems. An example of such ambiguity could be represented by the particular case of Wassily Kandinsky's painting *The Blue Rider* which gave its name to the homonymous art movement: it must be part of both artistic movement and painting categories, but at that level of thesaurus cataloguing no property value could distinguish one concept from the other. Another similar and frequent event is the case in which a work of art is simply called *Untitled* and only a highly structured set of related concepts (such as author, date of creation, etc.) could uniquely identify it among its homonyms. In conclusion, the strictness of the monohierarchical structure resulted unable to describe the complex relations linking art domain concepts. Nevertheless, the thesaurus has provided part of data and images and a first division plan for the ontology design.

## 3. Building the ontology

The ontology has been built with Protégé 3.1.1[3], one of the most popular ontology-editing environments. This release has been appreciably improved and it offers the opportunity to include pictures and hyperlinks that are embedded in instance editor forms. This new feature has been helpful to the purpose of our ontology, which has turn out enriched by visual content such as the reproduction of paintings in digital format taken by web sites and portals of art galleries and museums. Hyperlinks have been utilized to illustrate the predominant or descriptive colours of art works and colouring materials creating and loading monochrome html pages reproducing colour hues.

A fundamental source of information has been the WordNet 2.1 Browser lexical database[4] on which the central structure of the ontology has been based.

Starting from the WordNet top-level category called `entity` and descending along the major tree nodes, five main classes have been individuated (Figure 3):

- `Art_Movement`
- `Artist`
- `Art`
- `Colouring_Material`
- `Colour`

Subsequently, these classes have been deeply interrelated with each other by means of their own property slots forming a grid of ontology inner references (Figure 1). These relation slots constitute a major distinctive feature between our work and WordNet. In fact, the Art and Colours Ontology is mainly focused on instance and instance property structure, while WordNet, very similarly to the AAT or the ISO 21127, has been conceived as a framework for cataloguing purposes.

`Art` class hosts subclasses like `Plastic_Art` and `Ceramic` including the artisanship field in the art domain. Another important art subclass is `Graphic_Art` under which the `Painting` subclass has been included containing most of the sample instances. `Painting` class holds a further division of its content that has been structured in five provisional subclasses to which more could be added depending on the editor's classification needs: `Abstraction`, `Cityscape`, `Genre`, `Landscape` and `Monochrome`. Every single `Art` instance is described by several slots: *title*, *synonyms* (used to add one or more alternative art work titles), *date*, *authors* (`Painter`'s *name* is its inverse slot), *related_colouring_materials* (allowing multiple cardinality, it can be matched with pigments, paint and/or dyes) and *URL*. The URLwidget loads pictures or whole sites inside the form space or directly commands the default browser to open the Internet page.

The other potentially wide term repository is the `Artist` class. Its division is very similar to the `Art` class' although every `Artist` instance could be part of several classes, in case the artist had chosen to express himself through several artistic means. `Artist` class is described by *name* (the artist's surname), *synonyms* (formed by both forename and surname for a more flexible data retrieval, eventually more than one entry could be entered in presence of pseudonyms), *definition*, *place_of_birth*, *date_of_birth* and *date_of_death*. Moreover, every `Artist` instance is reciprocally related to *art_movement* and *art_works* slots.

`Colour` class has been divided into `Chromatic_Colour` and `Achromatic_Colour`. They are described by *name*, *synonyms* (used to assign different names to the same colour hue), *definition*, *RGB* (the red-green-blue triplet is used to describe digital format colours), *exadecimal_triplet* (the correspondence with the RGB value has been tested on Internet Explorer 6.0 and Konqueror 3.3 Internet browsers), *colouring_materials*

---

[2] http://www.niso.org/international/SC4/n500.pdf
[3] http://protege.stanford.edu
[4] http://wordnet.princeton.edu/perl/webwn

(`Colouring_Material`'s *name* is its inverse slot) and *related_art _works* (linked with one or more *title* slot of `Art` instances).

Plenty space has been given to `Colouring_Material` class, focusing on `Pigment` subclass and the *chemical_compound* slot. In order to access to on line resources, an *URL* slot has been provided.
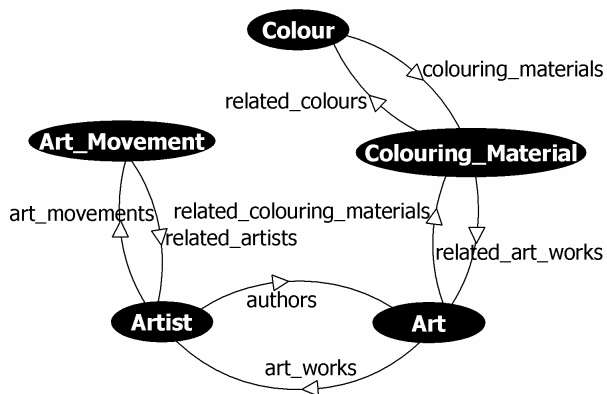
Moreover, the entire set of terms can be translated in other languages than English: *French_translation* and *Italian_translation* slots have been added to every class, although the translating possibilities are potentially unlimited.

Protégé 3.1.1 gives the possibility to question the ontology using a query editor. Choosing among the classes and instances composing the ontology and filling in the query masks, the user can access the knowledge base and display the result set with the opportunity to export the slot values to a text file. The query itself can be added to a query library for future references and an easy information retrieval independently from the stored data. Moreover, using the stored queries of the library, a further restricting condition can be imposed to the result set of the main query having the possibility to create complex structures of nested subqueries (Figure 2).

The whole ontology structure with its sample set of terms can be exported and browsed in various formats: pprj – the default Protégé project file, clips, html, rdf and owl.
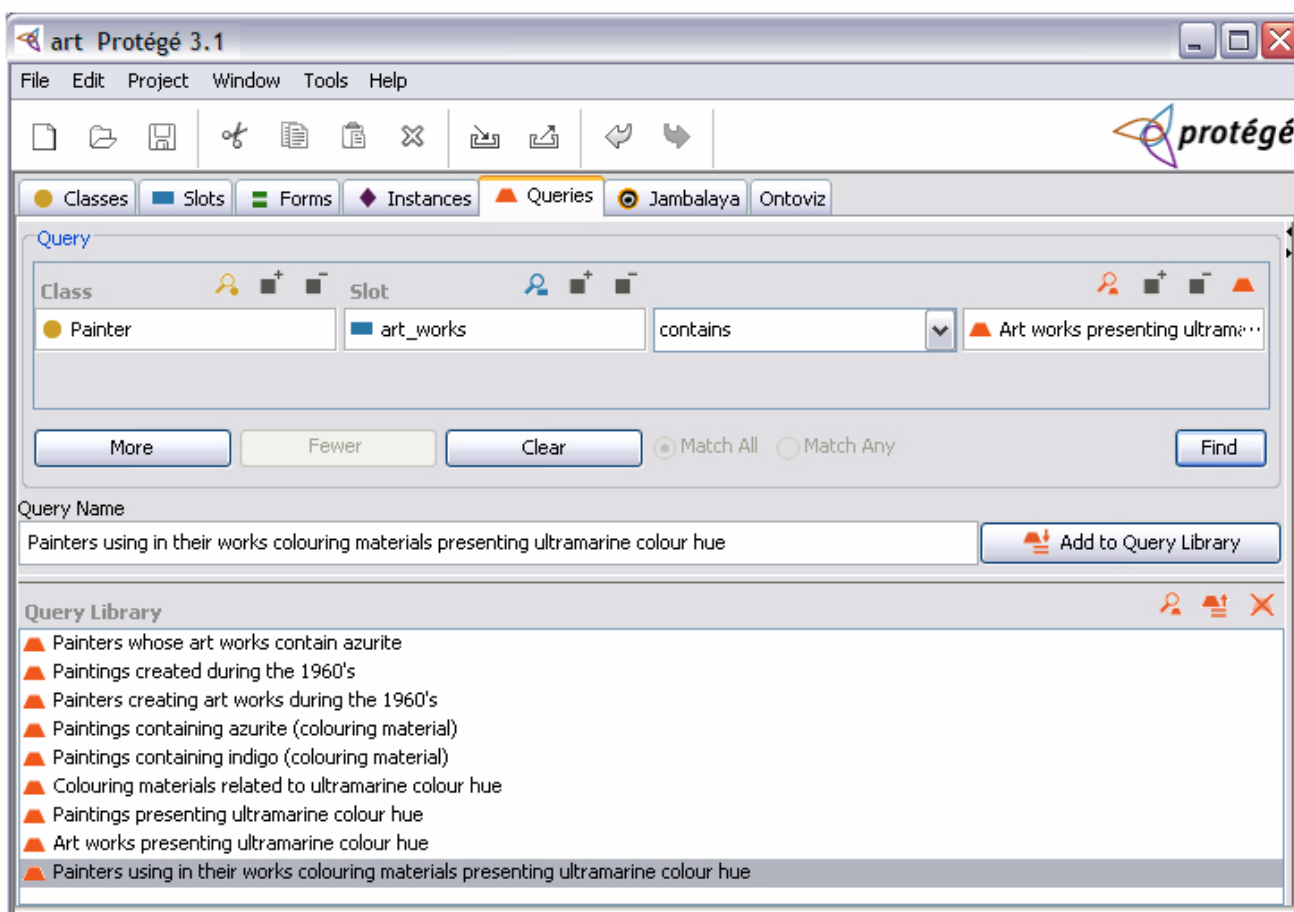


Figure 1: Art and Colours Ontology inverse slots linking classes.



Figure 2: Protégé 3.1.1 Query Library.

## 4. Conclusion

Building ontologies for large domains, such as medicine or art, is a costly affair. However, many domain thesauri have been built that can be a starting point, as in the present case study, for the construction of an ontology. The Art and Colours Ontology prototype could have potential applications in all those fields in which a classification of art works could help the information retrieval like museum web sites, art portals or universities, while the colouring material feature could respond to the requests of industry and chemistry for a rigorous pigment categorization. Furthermore, the ontology could be used as a knowledge base for educational purposes and as a lithe tool for researchers, scholars, common web users and artists themselves.

We plan to engage with the community of art and colours specialists to improve the results obtained with the present work.
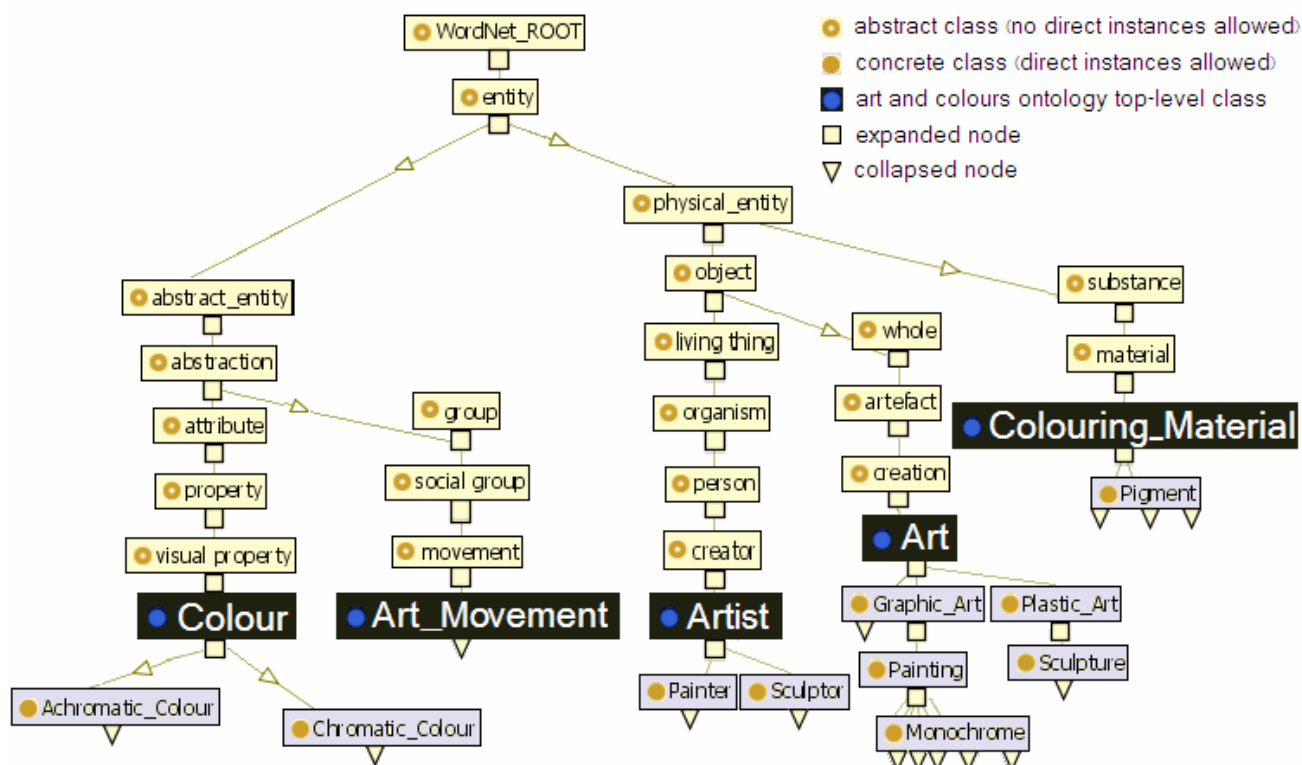
Figure 3: Hierarchical view of the Art and Colours Ontology classes.

## 5. References

Alani, H. (2001). Spatial and Thematic Ontology in Cultural Heritage Information Systems, Ph.D. dissertation. Computer Studies Department University of Glamorgan, U.K.

Fensel, D. (2001). *Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce,* Springer-Verlag.

Fernández-Lopez, M. & Gomez-Pérez, A. (2002). Overview and analysis of methodologies for building ontologies. *The Knowledge Engineering Review*, 17(2), pp.129-156.

Jaimes, A. & Smith, J.R. (2003). Semi-automatic, data-driven construction of multimedia ontologies. In *Proceedings IEEE International Conference On Multimedia and Expo*, Baltimore, MD, USA.

Kim, S., Alani, H., Hall, W., Lewis, P.H., Millard, D.E., Shadbolt, N.R., Weal, M.J. (2002). Artequakt: Generating Tailored Biographies with Automatically Annotated Fragments from the Web. In *Proceedings of the Semantic Authoring, Annotation and Knowledge Markup Workshop in the 15th European Conference on Artificial Intelligence,* Lyon, France, IOS Press.

Lefort, L. & Taylor, K. (2005). Large scale colour ontology generation with XO. In T. Meyer & M. Orgun (Eds.), *Conferences in Research and Practice in Information Technology*, Australian Computer Society, Inc.

Peterson, T. (1994). *Introduction to the Art and Architecture Thesaurus*. Oxford University Press.

Wielinga, B.J., Schreiber, A.Th., Wielemaker, J., Sandberg, J.A.C. (2001). From Thesaurus to Ontology. In Y. Gil, M. Musen, J. Shavlik (Eds.), *Proceedings 1st International Conference on Knowledge Capture*, Victoria, Canada, pp. 194-201, ACM Press.