

## WP8 Case study: Cultural Heritage

3rd MOLTO Project Meeting  
Helsinki, 31 August 2011

- To build an ontology-based multilingual grammar for museum information starting from the CIDOC-CRM ontology for artefacts at Gothenburg City Museum
- To cover 15 languages for baseline functionality and 5 languages with a more complete coverage
- To build a prototype of a cross-language retrieval and representation system to be tested with objects in the museum, and automatically generate Wikipedia articles for museum artefacts in 5 languages

- D8.1 Ontology and corpus study of the cultural heritage domain (M18)
- D8.2 Multilingual grammar for museum object descriptions (M24)
- D8.3 Translation and retrieval system for museum object descriptions (M30)

- D8.1 Knowledge management software for museums databases
- D8.2 Natural language generation of descriptions of museum artefacts on mobile platforms  
multilingual catalogs (online or offline)
- D8.3 Multilingual editor for updating the museum artefact database

- Empirical study of existing metadata schemas adopted by museums in Sweden
- Ontology development supporting compatibility to a variety of CH data schemata
- Analysis of syntactic structures and patterns for discourse generation
- Verbalization of the ontologies in GF ( language generation prototype for English and Swedish )

## Record fields found in the Gothenburg City Museum database

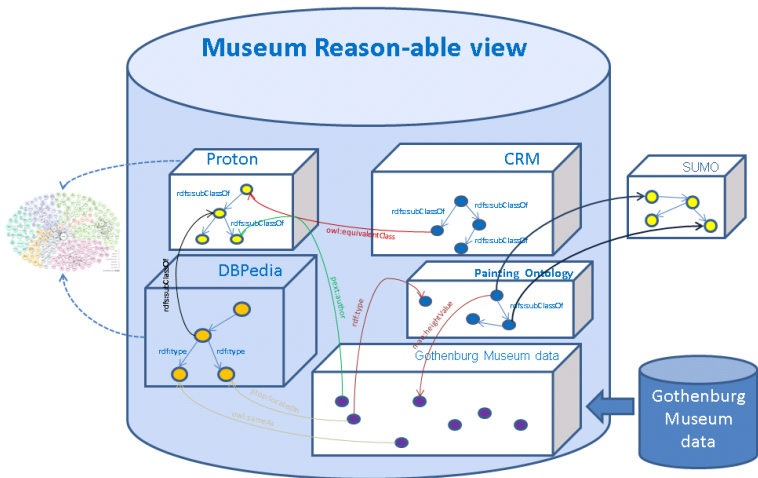
<b>Database record field</b>	<b>Value</b>
Style	Baroque
Technique	Glaze
Medium	Linseed oil
Color	Gold
Category	Western Painting
Provenience (object originates)	Norrköping
Price/Value	3000 kr
Association number	GIM8165
Association Date	940308
Method of acquisition	Present

Purpose: to support integration and interoperability of the CIDOC-CRM ontology with other ontologies and schemata, including:

- CIDOC-CRM: describing the implicit and explicit concepts and relationships used in cultural heritage documentation
- SUMO: Merge and Mid-Level Ontology
- Swedish Open Cultural Heritage (SOCH) provides a metadata that helps to intermediate data between museums in Sweden and the Europeana portal

- The painting ontology is developed in Protégé
- Its main reference model is the OWL 2 implementation of the CRM.
- Integration of concepts is accomplished by using the OWL construct *intersectionOf*
- 197 classes  
24 stems from CRM, 15 equivalent to SOCH, 45 equivalent to SUMO concepts
- 107 properties  
17 are subproperties of the CRM properties





*[TITLE CREATOR DATE SIZE COLOR LOCATION]*

Guernica was painted by Pablo Picasso in 1937. It is of size 349 on 776 cm and is painted in white, black and gray. The painting is displayed at Museo Reina Sofia from Madrid.

*[TITLE TYPE CREATOR LOCATION DATE]*

Guernica is an oil painting by the Spanish artist Pablo Picasso. The painting was displayed at Museo Reina Sofia in 1937.

*[TITLE TYPE LOCATION CREATOR SIZE]*

The Last Supper is a mural painting in Milan created by Leonardo da Vinci. It measures 450 on 870 centimeters.

*[TITLE TYPE CREATOR PLACE DATE LOCATION]*

The Massacre at Chios is an oil painting by the French artist Eugène Delacroix. The painting was completed and displayed at the Salon of 1824 and presently hangs at the Musée du Louvre in Paris.

*[TITLE TYPE DATE CREATOR]*

Valdemar Atterdag holding Visby to ransom is an oil on canvas, signed in 1882 by the Swedish historical painter Carl Gustaf Hellqvist (1851 – 1890).

*[TITLE TYPE CREATOR DATE SIZE]*

Olympia is an oil on canvas painting by Édouard Manet. Painted in 1863, it measures 130.5 by 190 centimetres (51 x 74.8 in).

- PN -> NP  
Van Gogh
- Det -> CN -> NP  
The portrait  
The countess of Carnarvon
- NP -> Adv -> NP  
The bell in London
- V2 -> PP -> VP  
displayed at the Paris Salon  
painted by Jamie Wyeth
- V2 -> Adv -> VP  
displayed here suggest the hand of an artist
- V2 -> NP -> VP  
displays painting of tulip bearing her signature

Straightforward from the ontology:

- Guernica is a painting.
- Guernica is painted on canvas.
- Guernica is painted by Pablo Picasso.
- Guernica was created in 1937.

# Generation grammar adapted for discourse rendering

Discourse pattern:

[*TITLE CREATOR DATE SIZE COLOR LOCATION*]

Guernica was painted by Pablo Picasso in 1937. It is of size 349 on 776 cm and is painted in white, black and gray. The painting is displayed at Museo Reina Sofia from Madrid.

Guernica blev målad av Pablo Picasso år 1937. Den är av storlek 349 och 776 cm och är målad i vit , svart och grå. Målningen återfinns på Museo Reina Sofia i Madrid.

- Implement more patterns for discourse generation
- Translate lexical entites and write grammar for French, Finnish and German
- Map SPARQL query results to GF abstract syntax (WP4)

- Which lexical entries to translate?
- How many descriptions should we aim at?
- The XML data specification from SPARQL to GF
- How to evaluate the output from GF?



- Damova, M., Dannélls D. (2011). Reasonable View of Linked Data for Cultural Heritage. The Third International Conference on Software, Services & Semantic Technologies (S3T)
- Dannélls, D., Damova M., Enache R., Chechev M. (2011). A Framework for Improved Access to Museum Databases in the Semantic Web. Language Technologies for Digital Humanities and Cultural Heritage. In conjunction with Recent Advances in Natural Language Processing (RANLP)
- Dannélls D. The painting ontology. Journal of applied ontologies (Submitted)