

Online multilingual generation of Cultural Heritage content

Dana Dannélls

Språkbanken, Department of Swedish Language
University of Gothenburg

MOLTO 2012
2012-03-07

Motivation

- ▶ New developments in technologies (e.g. Semantic Web) provide sophisticated information access to cultural heritage material
 - ▶ e.g. enables users to broaden/narrow search based on multiple criteria at once
- ▶ Emerging European project initiatives provide cross-collection, cross-museum and cross-subject access to bigger sets of data collections
 - ▶ MultimediaN E-culture, Europeana, Cornucopia, Michael

Direct access to cultural heritage objects

TOCET VIEWS ▾

Who	What	Where
Modigliani, Amedeo	140 papier	28483 Rijksmuseum_Volkenkunde
Cézanne, Paul	127 prent	23525 BiblioPolis
Picasso, Pablo	112 handwriting (1)	
David, Jacques-Louis	95 foto	
Rembrandt van Rijn	91 Foto	
Degas	n/a	

Work What: oil paint

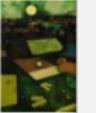
 Printer and inker in...
Venne, Adriaen van de

 Compositor in the pr...
Venne, Adriaen van de

 The Van de Venne
Venne, Adriaen van de

 Ein Karem

 Triptych to the Fallen

 For the Fallen

The Van de Venne brothers in their printing shop, together with Apollonius S

<http://e-culture.multimedian.nl/ns/biblioPolis#1497>

An author (with glasses) is busy correcting a manuscript in the print shop. Behind him Adriaen en Jan van de Venne.

links

- full view
- annotate

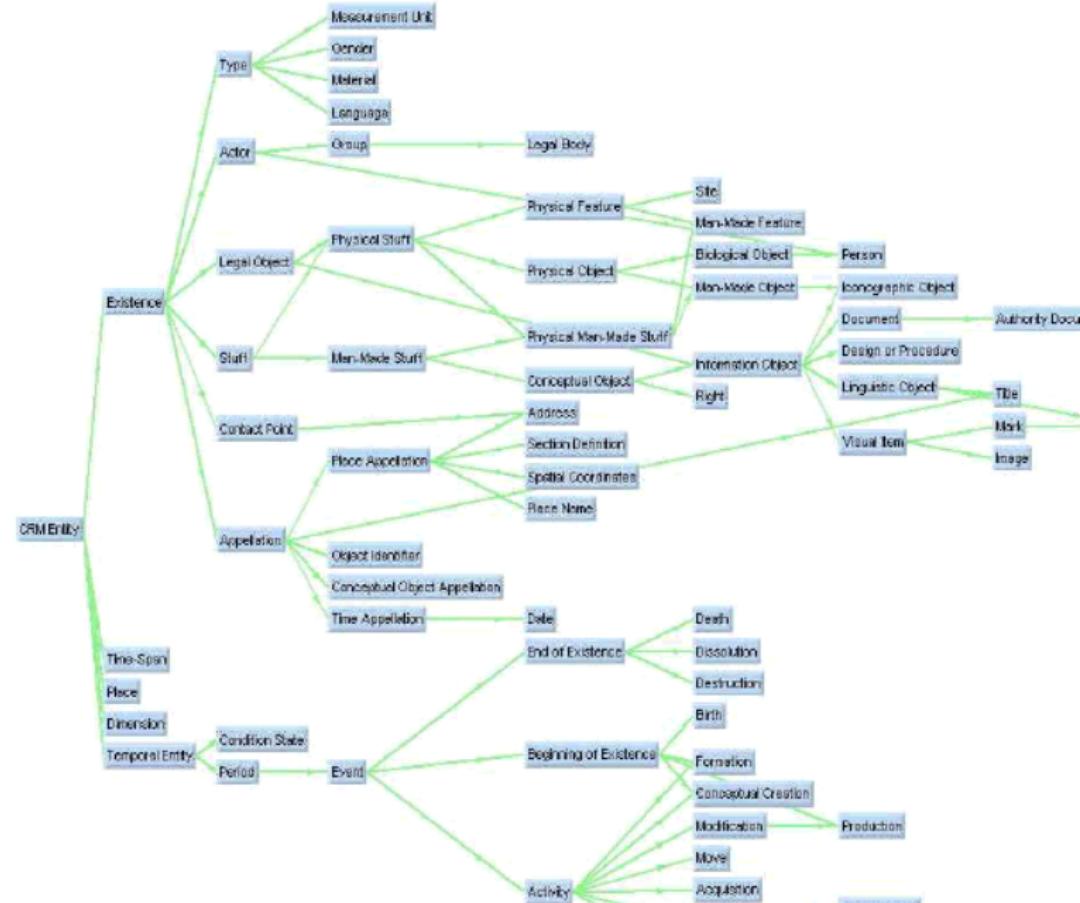
Property	Value
Creator	Venne, Adriaen van de
Date	1623
Location	BiblioPolis; Daily Mail and General
Measurements	46,5 x 76,2 cm.
Subject	Geschiedenis van de boekdrukku authors; printing houses; proof re
Technique	oil paint; olieverf
Title	The Van de Venne brothers in the shop, together with Apollonius S gebroeders Van de Venne in hun tesamen met Apollonius Schotte
Type	easel paintings; schilderij

The CIDOC Conceptual Reference Model (CIDOC-CRM)

The CIDOC Conceptual Reference Model (CIDOC CRM), developed by the International Committee for Documentation (CIDOC) of the International Council of Museums (ICOM) (Crofts et al., 2008).

- ▶ ISO standard since 2006
- ▶ 87 classes and 130 relationships
- ▶ Available in OWL

The CIDOC Conceptual Reference Model (CIDOC-CRM)



Project goals

- ▶ 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

A record from the Gothenburg City Museum database

Field name	Value
Field nr.	4063
Prefix	GIM
Object nr.	8364
Search word	painting
Class 1	353532
Class 2	Gothenburg portrait
Amount	1
Producer	E.Glud
Produced year	1984
Length cm	106
Width cm	78
Description	oilpainting represents a studio indoors
History	Up to 1986 belonged to Datema AB, Flöjelbergsg 8, Gbg
Material	oil colour
Current keeper	2
Location	Polstjärnegatan 4
Package nr.	299
Registration date	19930831
Signature	BI
Search field	BO:BU Bilder:TAVLOR PICT:GIM

The Painting ontology

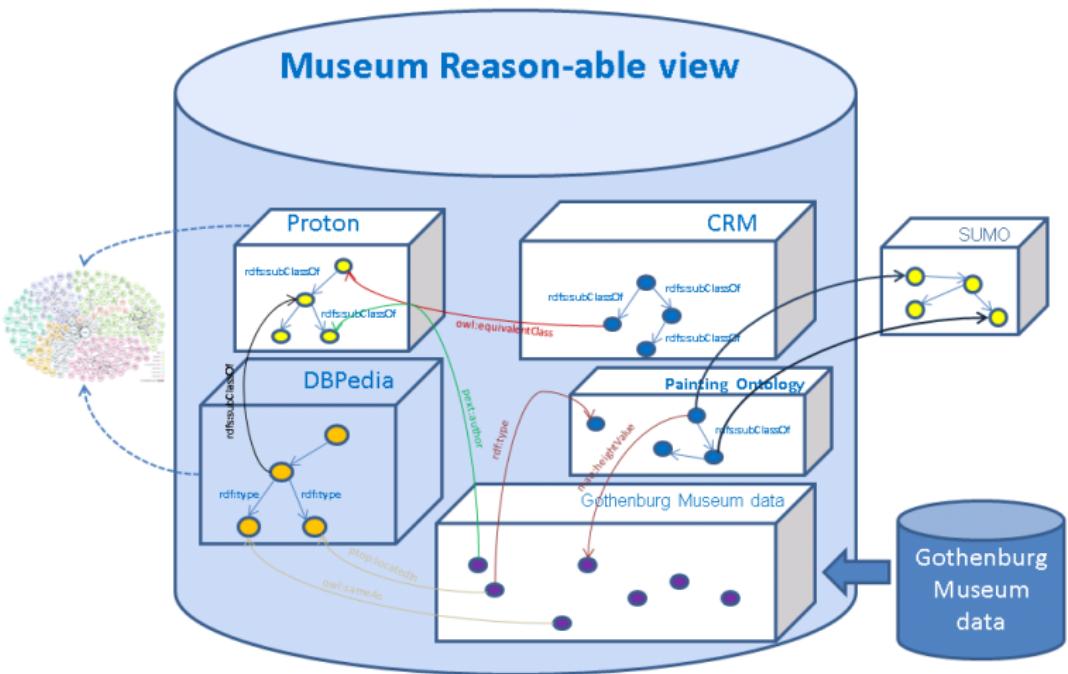
Purpose: to support integration and interoperability of the CIDOC-CRM ontology with other ontologies and schemata, including:

- ▶ CIDOC-CRM
- ▶ SUMO: Merge and Mid-Level Ontology
- ▶ Swedish Open Cultural Heritage (SOCH)

The painting ontology contains:

- ▶ 197 classes
24 stems from CRM, 15 equivalent to SOCH, 45 equivalent to SUMO concepts
- ▶ 107 properties
17 are subproperties of the CRM properties

Integration of Gothenburg City Museum data



Museum Reason-able View Environment

GIM8165Obj

Source: <http://sprakbanken.gu.se/rdf/owl/painting.owl#GIM8165Obj>

Subject (22) Predicate Object All

Statements in which the resource exists as a subject.

Predicate	Object
rdf:type	crm:E24_Physical_Man-Made_Thing painting:Painting painting:tavia_diplom
crm:P43_has_dimension	painting:GIM8165Quantity
painting:acquiredThrough	painting:GIM8165Acquisition
painting:belongsTo	painting:GiM
painting:createdBy	painting:Lith_Akt_Bol_Norrköping
painting:hasClassification	painting:FoodIndustry
painting:hasCreationDate	painting:GIM8165CreationDate
painting:hasCurrentLocation	painting:Postgårdsgatan_4
painting:hasDimension	painting:GIM8165ImageDimension
painting:hasHistoricalContent	painting:GIM8165Description painting:GIM8165History
painting:hasMaterial	painting:Wood
painting:hasRepresentation	painting:Bo.BU_Bilder:LIVSMEDELSIND_PICT.GIM_8165

8 thousand museum artifacts from the Gothenburg city museum database.

Ontology verbalization in GF

Straightforward from the ontology:

- ▶ isA (Object, Painting)
Guernica is a painting.
- ▶ createdBy (Painting, Creator)
Guernica is created by Pablo Picasso.
- ▶ hasCreationDate (Painting, TimeSpan)
Guernica was created in 1937.

A case study on English and Swedish

The corpus data for analysis

- ▶ 40 parallel texts extracted from Wikipedia under the category *Painting*
- ▶ 300 object descriptions for each language extracted from museums online databases

The results of the analysis

- ▶ a list of syntactic structures
- ▶ a list of discourse patterns

Syntactic structures

- ▶ 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

Discourse patterns

- ▶ DP0 : painting painter year -> Text
- ▶ DP1 : painting museum painter size -> Text
- ▶ DP2 : painting painter represented museum -> Text
- ▶ DP3 : painting material year painter -> Text
- ▶ DP4 : painting painter year museum colour size -> Text

Discourse patterns generation in GF I

DP0

- ▶ (eng) Sommer Joy was painted by Anders Zorn.
- ▶ (swe) Sommarnöje blev målad av Anders Zorn.

DP1

- ▶ (eng) Sommer Joy was painted in 1886. It measures 349 by 776 cm.
- ▶ (swe) Sommarnöje blev målad år 1886. Den är av storlek 349 och 776 cm.

DP3

- ▶ (eng) Sommer Joy is painted on paper in 1886 by Anders Zorn.
- ▶ (swe) Sommarnöje blev målad på papper 1886 av Anders Zorn.

Discourse patterns generation in GF II

DP2

- ▶ (eng) Sommer Joy is a painting made by Anders Zorn. The work depicts a view from Lilla Bommen at Hisingen.
- ▶ (swe) Sommarnöje är en målning av Anders Zorn. Den föreställer en utsikt från Lilla Bommen mot Hisingen.

DP4

- ▶ (eng) Sommer Joy was painted by Anders Zorn in the year 1886. It is of size 349 by 776 cm and is painted on paper. The painting is displayed at the Museum of World Culture.
- ▶ (swe) Sommarnöje blev målad av Anders Zorn år 1865. Den är av storlek 349 och 776 cm och är målad på paper. Målningen återfinns på Världskulturmuseet.

A description of a museum object

MOLTO is funded by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement FP7-ICT-247914.

GSM940051Obj

Hisingen was painted by Brynolf Wennerberg in 1889. Hisingen blev målad av Brynolf Wennerberg år 1889. Hisingen is a portrait made by Brynolf Wennerberg in 1889. The portrait is displayed at the Gothenburg City Museum. Hisingen är en porträttmålning av Brynolf Wennerberg från 1889 som finns utställd i Göteborgs stadmuseum. Hisingen är en målning gjord av Brynolf Wennerberg 1889. Den föreställer Oljemålning porträtt om en man med hatt och vitkravat. Målningen är målad på trä. Dimensionerna är 349x776 cm. Hisingen är en målning som finns i Göteborgs stadmuseum. Den är målad i olja på träplatta och är från 1889.

URI: <http://tegelskatten.gu.se/triplestore/painting.owl#GSM940051Obj>

Subject (36) Predicate Object All

Named Graph All Language

Statements in which the resource exists as a subject.

Predicate	Object
rdfs:label	Gsm E24_Physical_Man-Made_Thing Painting_OilPainting Painting_Painting Painting_mining owl:NamedIndividual
cm:has_dimension	Painting_GSM940051Quantity
painting_acquiredThrough	Painting_GSM940051Acquisition
painting_belongsTo	Painting_GSM
painting_createdBy	Painting_BrynolfWennerberg Painting_Brynolf_Wennerberg
painting_hasAcquisitionDate	Painting_GSM940051AcquisitionDate
painting_hasAssignedAttribu...	Painting_Good
painting_hasClassification	Painting_GothenburgMotive Painting_Paintings
painting_hasCreationDate	Painting_GSM940051CreationDate

Current state of work

- ▶ Implementing more patterns for discourse generation
- ▶ Building a lexicon to cover the content of all object descriptions
- ▶ Translate lexical entities and write grammar for Finnish, French and German