

WP8 Case study: Cultural Heritage

Dana Dannélls, Aarne Ranta, Ramona Enache,
Mariana Damova, Maria Mateva

University of Gothenburg and Ontotext

August 2013

The aim of this work

To build an ontology-based application for communication of museum content on the Semantic Web and make it accessible in 15 languages.

Outline

- Museum Reason-able View: Interoperable cultural heritage knowledge bases
- Ontology-based multilingual grammar for retrieving and generating museum content:
 - RDF to NL – well-formed descriptions
 - NL to RDF – SPARQL queries linearization
- Cross-language retrieval and representation system using Semantic Web technology

The Museum Reason-able View (MRV)

Domain ontology: CIDOC Conceptual Reference Model (CRM) v. 5.0.1

Application ontologies: Painting ontology and Museum Artifacts Ontology (MAO)

Ontology	Classes	Properties
CIDOC-CRM	87	130
Painting ontology	197	107
MAO	10	20
Total	836	440

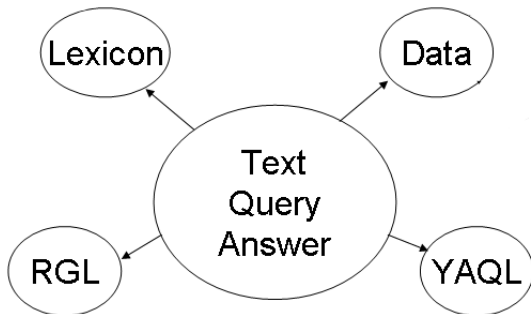
The museum Linked Open Data (LOD)

Gothenburg City Museum (GCM): Only two collections (GSM and GIM)

DBpedia: Larger amount of painting entities

Source	Amount of entities
GCM	48
DBpedia	614
Total	662

The application grammar overview



Supported languages

Bulgarian, Catalan, Danish, Dutch, English,
Finnish, French, Hebrew, Italian, German,
Norwegian, Romanian, Russian, Spanish, Swedish.

The lexicon grammar

Covers a subset of the ontology classes and properties

Classes and properties were manually translated

Input:

createdBy (Ophelia, Brynolf Wennerberg)

isA (Brynolf Wennerberg, **Painter**).

isA (Ophelia, **Painting**)

Direct verbalization:

Ophelia is a painting. Ophelia was created by Brynolf Wennerberg. Brynolf Wennerberg is a painter.

Our approach:

Ophelia was painted by Brynolf Wennerberg.

The data grammar

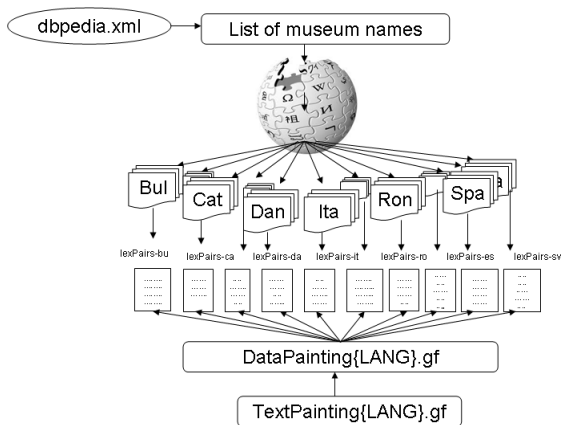
Contains ontology entities that were extracted from GCM and DBpedia

No adequate translations from DBpedia

Entities of museum names are automatically translated from Wikipedia

Class	Entities
Title	662
Painter	116
Museum	104
Place	22
Total	904

Automatic translation process overview



Available museum names translations

Language	Translated (out of 104)
Bulgarian	26
Catalan	63
Danish	33
Dutch	81
Finnish	40
French	94
Hebrew	46
Italian	94
German	99
Norwegian	50
Romanian	27
Russian	87
Spanish	89
Swedish	58

Text grammar

Nine classes: Title, Painter, Type, Colour, Size, Year, Material, Museum, Place

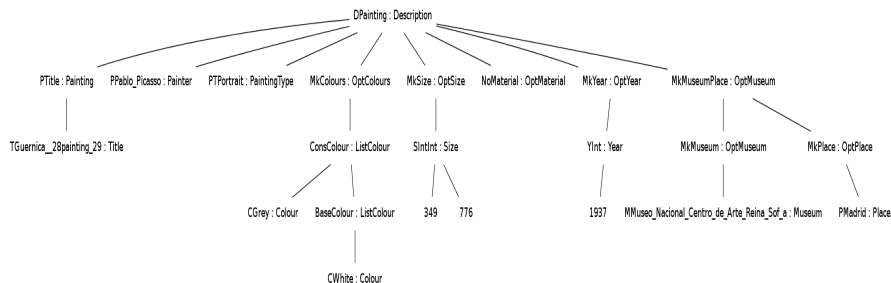
Three are obligatory

Interior was painted by Edgar Degas.

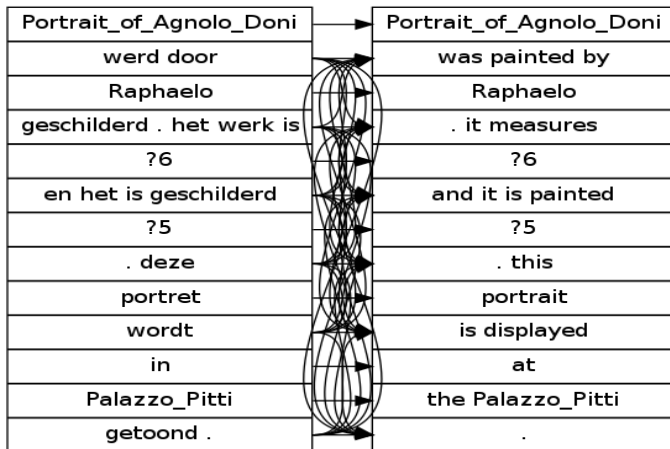
One function, three sentences

Interior was painted on canvas by Edgar Degas in 1868. It measures 81 by 114 cm and it is painted in red and white. This painting is displayed at the Philadelphia Museum of Art.

Text grammar: RDF to NL



Word alignment example



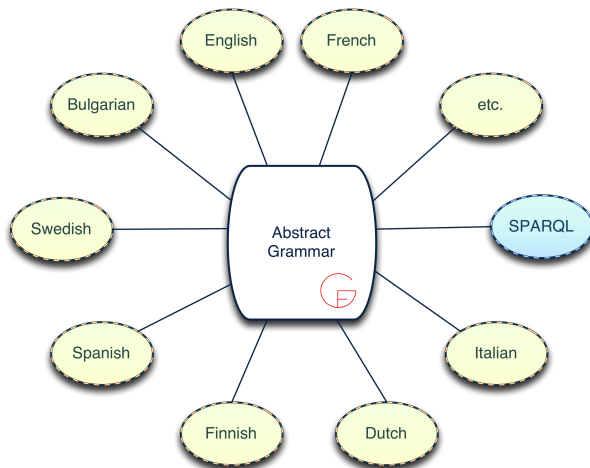
Multilingual challenges

- Lexicalizations
Classes: compounds, multiword expressions
Properties: verbs, adverbs, prepositions
- Order of semantic elements
Material, Year
- Tense and voice
Past, past participle, present, active/passive
- Aggregation
Conjunction, relative clause, punctuations
- Coreference
Pronoun, noun, empty reference

Multilingual examples

TextPainting: DPainting (PTitle TInterior__28Degas_29) PEdgar_Degas PTPainting NoColours (MkSize (SIntInt 50 103)) NoMaterial (MkYear (YInt
TextPaintingBul: Interior е нарисувана от Edgar Degas през 1890 година. Тя е с размер 50 см на 103 см. Този експонат е изложен в Лувър.
TextPaintingCat: Interior fou pintat per Edgar Degas en 1890. Mesura 50 per 103 cm. Aquesta pintura està exposada al Museu del Louvre.
TextPaintingDan: Interior blev malet af Edgar Degas i 1890. Det er 50 ganger 103 cm. Dette maleri er udstillet på Louvre.
TextPaintingDut: Interior werd in 1890 door Edgar Degas geschilderd. Het werk is 50 bij 103 cm. Dit schilderij wordt in Musée du Louvre get
TextPaintingEng: Interior was painted by Edgar Degas in 1890. It measures 50 by 103 cm. This painting is displayed at the Musée du Louvre.
TextPaintingFin: maalauksen Interior on maalannut Edgar Degas vuonna 1890. Se on kokoa 50 kertaa 103 cm. Tämä maalaus on esillä Louvressa.
TextPaintingFre: Interior a été peint par Edgar Degas en 1890. Il est de 50 sur 103 cm. Ce tableau est exposé au Musée du Louvre.
TextPaintingGer: Interior wurde in 1890 von Edgar Degas gemalt. Das Werk ist 50 mal 103 cm. Dieses Bild ist ausgestellt im Der Louvre .
TextPaintingHeb: רבועה | מאיזומב תגצמ וז הרצי. רתמ טנס 103 לע 50 ול וגב אה. 1890 תגצמ סגד לר וצ אה סיפ.
TextPaintingIta: Interior è dipinto da Edgar Degas in 1890. Misura di 50 su 103 cm. Questo dipinto è esposto al Museo del Louvre.
TextPaintingNor: Interior ble malt av Edgar Degas i 1890. Det er 50 ganger 103 cm. Denne malerien er utstilt på Musée du Louvre.
TextPaintingRon: Interior este pictat de către Edgar Degas în 1890. Este din 50 pe 103 cm. Acest tablou este expus în Musée du Louvre.
TextPaintingRus: Interior нарисовалось Edgar Degas в 1890 году. Она с размером 50 см на 103 см. Эта картина видится в Лувр.
TextPaintingSpa: Interior fue pintado por Edgar Degas en 1890. Mide 50 por 103 cm. Esta pintura está expuesta en el Museo del Louvre.
TextPaintingSwe: Interior målades av Edgar Degas år 1890. Den är 50 gånger 103 cm. Den här målningen är utställd på Louvren.

Query grammar: NL to SPARQL



Grammar advantage/disadvantages

- + Modular grammar design
- + The lexicon and the data are shared by all grammars
- + 16 text patterns from one function
- + Natural realization of the ontology content
- + Takes 30 minutes to implement a new language if the language is in the RGL
- The texts can become artificial if the names are missing translations

Grammar writing

It may take two to five days to add a new language if

- the language is not in the RGL (e.g. Hebrew)
- the grammarian is familiar with GF
- and he/she acquires good knowledge of the language

Demo

<http://museum.ontotext.com/>