

STOLE: A Reference Ontology for Historical Research Documents

Laura Pandolfo

Ph.D. in Digital Humanities
DIBRIS – University of Genoa

`laura.pandolfo@edu.unige.it`



AI*IA 2015
Doctoral Consortium
23 September 2015, Ferrara

Digitization of Cultural Heritage materials

In the last years, Cultural Heritage (CH) institutions have promoted numerous projects about the digitization of their cultural objects.

→ Applications and tools that allow to get access to digital CH.

- Europeana is an example of one of the most recent big such efforts.

Frequently, these projects involve researchers coming from different areas, such as Humanities and Knowledge Engineering.

Motivation (1/2)

Benefits of CH digitization:

- Contribute to the conservation and preservation of heritage.
- Create new educational opportunities and encourage tourism.
- Provide new ways of improving access to this patrimony.

CH represents a complex domain characterised by a multidisciplinary approach.

- Need effective solutions about the design and the implementation of user-friendly ways to access and query content.

Motivation (2/2)

Semantic Web technologies meets Cultural Heritage

This application domain has become one of the most successful fertile ground of Semantic Web technologies because of its own special features.

- Data are syntactically and semantically heterogeneous, multilingual and highly interlinked.
- Cultural objects are produced in a distributed, open fashion by several different memory institutions.
- Managing publication of such rich and wide content on the web poses new challenges.

Our Contribution

STOLE Ontology

A reference domain ontology that provides a vocabulary of terms and relations with which it is possible to model the history of Italian public administration.

- It represents a first step towards the development of a knowledge exchange standard on this specific domain.
- To the best of our knowledge, this is the first time that ontological modeling has been undertaken for this specific domain.

Outline

- 1 Case Study
- 2 STOLE Ontology
- 3 Current & Future Work
 - Ontology Population
 - Modeling Temporal Changes of Entities

Case Study

Historical Research is a systematic process of searching for the facts and then using that information to describe, analyze, and interpret the past (W. Wiersma, 2000).

The History of the Italian Public Administration relates the understanding the history of institutional changes, with particular attention to the evolution of public administration in Italy.

One of the main sources used by historians are journal articles at the time.

- These documents represent a valuable source of information used by historians, political scientists and administrative law experts.

Outline

- 1 Case Study
- 2 STOLE Ontology**
- 3 Current & Future Work
 - Ontology Population
 - Modeling Temporal Changes of Entities

STOLE Ontology (1/3)

The main goal of STOLE is to model historical concepts and gain insights into this specific field in order to support historians in their research tasks.

- Historical analysis is based on specific categories of information and focused on the interrelations between these data.

For example, the relationship between an author and the people cited in an article could provide valuable information to historians.

STOLE Ontology (2/3)

It has been modeled by means of existing standards and meta-data vocabularies:

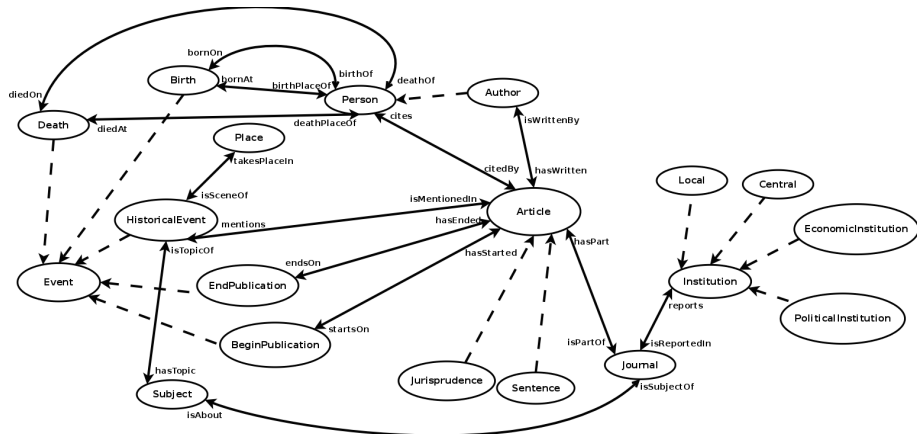
- Bibliographic Ontology
- Bio Vocabulary
- Dublin Core
- FOAF
- Ontology of the Chamber of Deputies (OCD)

The STOLE ontology has been developed using OWL 2 DL language.

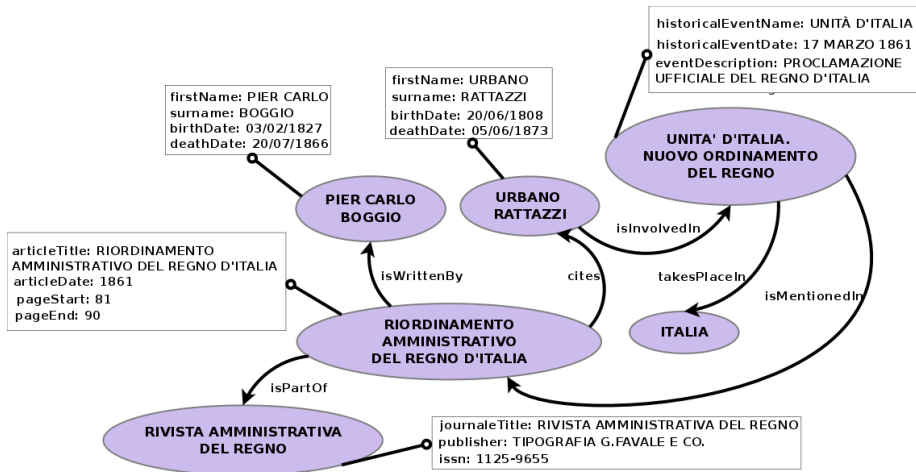
Axioms	465
Classes	19
Object Properties	34
Data Properties	33

Table : TBox Metrics

STOLE Ontology (3/3)



STOLE Ontology - Example



Outline

1 Case Study

2 STOLE Ontology

3 Current & Future Work

- Ontology Population
- Modeling Temporal Changes of Entities

Outline

- 1 Case Study
- 2 STOLE Ontology
- 3 Current & Future Work
 - Ontology Population
 - Modeling Temporal Changes of Entities

Current & Future Work

Ontology Population

The ontology population process from texts is very time-consuming and it requires specialized skills.

- Recent contributions about the automatization of this process are based on the following techniques:
 - ▶ Natural Language Processing.
 - ▶ Information Extraction.
 - ▶ Machine Learning.
- Once the ontology will be fully populated, we are planning to perform an experimental analysis on the STOLE ontology.

Outline

- 1 Case Study
- 2 STOLE Ontology
- 3 Current & Future Work
 - Ontology Population
 - Modeling Temporal Changes of Entities

Current & Future Work

Modeling Temporal Changes of Entities

We intend to model temporal transformation through Interval Temporal Description Logics, which provide a set of temporal operators.

- Institution that changed its functions over time.
 - ▶ The institution *Corte di Cassazione* was created in 1848 as a local institution, then it became a central institution in 1923, during Fascismo.
- Artale, A., Kontchakov, R., Ryzhikov, V., Zakharyashev, M.: Tractable interval temporal propositional and description logics. In: Proc. AAAI 2015.
- Artale, A., Kontchakov, R., Ryzhikov, V., Zakharyashev, M.: Interval Temporal Description Logics. DL 2015.

Thank you!