News Ontology

OVERVIEW

The purpose of the News Ontology is to build a model which allows the alignment across the diciplines of news professionals and historians.

This news ontology has been designed using the RDF formalism and is based upon several main classes which contain the "real world" objects we deal with. It proposes a model to describe the implicit knowledge previously extracted and annotated by Papyrus tools (data mining, classifiers) from the news corpora.

Once populated with news items, the ontology shows the relations between the documents, named entities (persons, organizations, locations, events, works) and Papyrus themes, which allows to navigate between news and concepts (abstract concepts and named entities). These relations are subsequently used by the History ontology to match Historian queries with pertinent news items.

INNOVATION

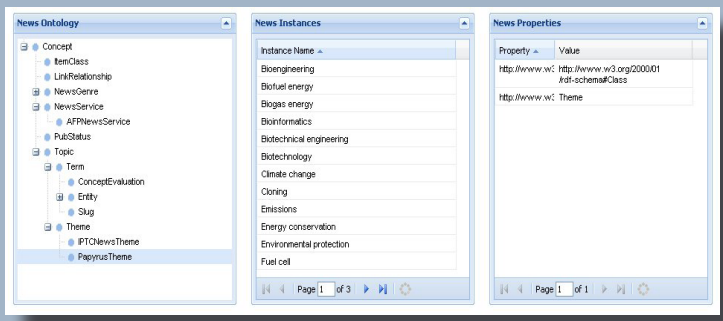
This News Ontology is the first attempt to build an OWL IPTC compliant model to describe the news items implicit knowledge and especially the relations between concepts, named entities and themes.

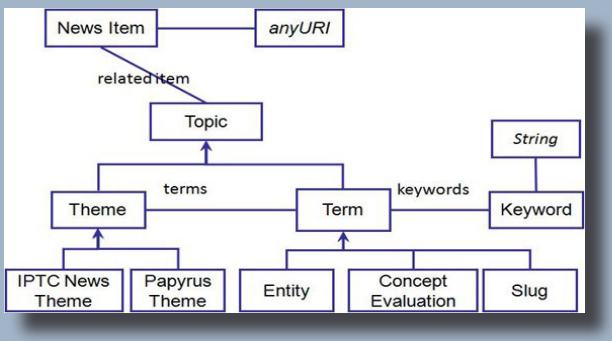
Previous attempts were limited to extract list of taxonomies without necessarily linking concepts between them.

Moreover, the main innovation of this project is to achieve a mapping between items from different sources (history and news) via the Papyrus themes which are transverse between the two domains.

BUSINESS IMPACT

This ontology, enhanced with the knowledge related to an application domain, could be used with archivist content management systems to





enable Historians to perform queries on archived news databases.

Papyrus theme's instances can be replaced by other domains related taxonomies to perform similar navigation into news corpora.

INTEROPERABILITY

Being designed in RDF, a standard of W3C, this ontology may be reused or enhanced with new classes of other domains. It can also be imported with the RDF/OWL import process within another ontology.

The Papyrus news ontology is available in open source through the project portal (www.ict-papyrus.eu).

Contact: Denis Teyssou - denis.teyssou@afp.com
PAPYRUS is a project partly funded by the EC FP7 ICT Programme