Exploration of Text–Object Relationships with Semantic Web

(with special reference to Arabic language)

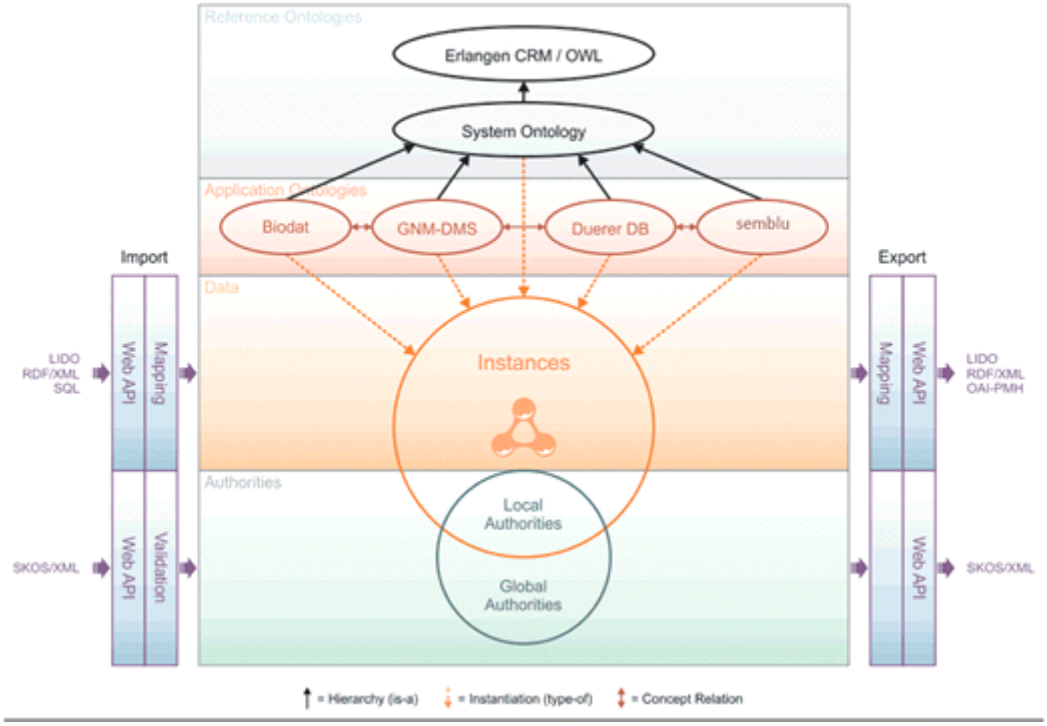
Sree Ganesh Thottempudi

Abstract

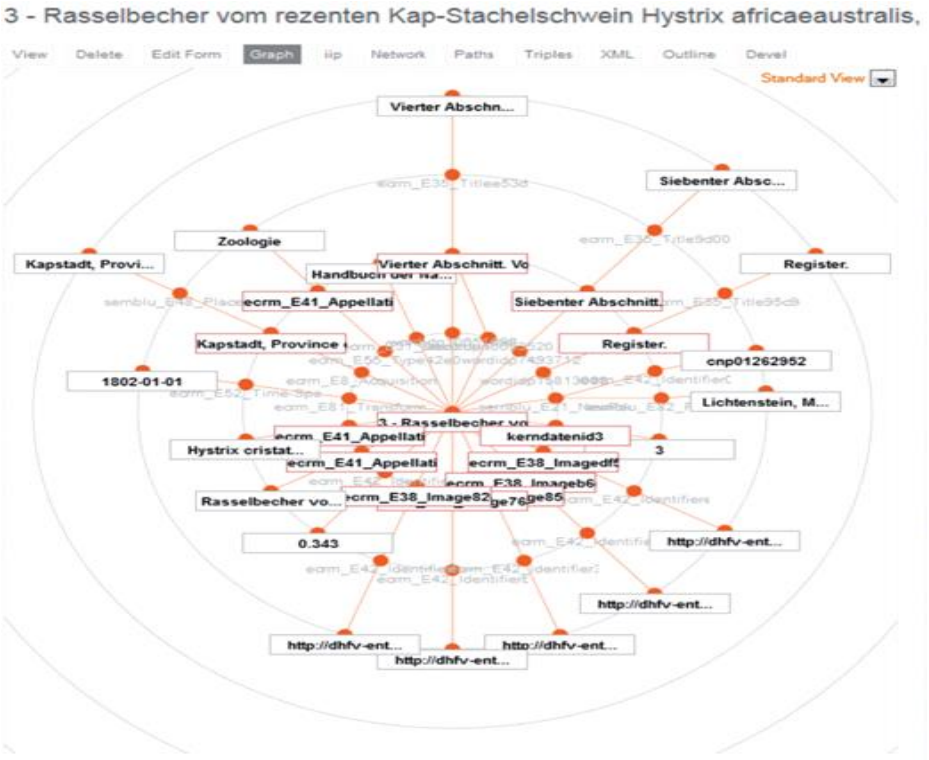
Keywords: TEI, Annotation, RDF, Ontology, Triplestore data, Semantic Web, Python

Better information management is the key to a more intelligent DH. every country has lot of information related to their heritage and culture in the form of manuscripts. To retrieve the science and heritage information from these manuscripts is time consuming and people dont have interest to through them. DH is putting its efforts to retrieve the scientific information (Chula A. Eide Ø. 2014). . In this direction, many challenges must be first overcome. Enabling seamless, effective and efficient access to the various data sets and novel methods for exploiting from ancient manuscripts. This workshop aims to bring together an interdisciplinary audience interested in the fields of semantic web, data management and ancient manuscripts to discuss the unique challenges in Text — Object relationship establishment (Eide Ø. 2013) We are doing hands on training for data management and to propose novel and practical solutions for the preservation of ancient heritage and culture to the next generation. As semantic technologies are currently widely exploited more and more for the management of DH data, new challenges occur while processing different kinds of ancient manuscripts data that dictate new solutions. Establishing Text and object relation is one of the challenges.

In this workshop we are giving basic training in Python. Students need not be a expert in programming. From that point we are giving training towards TEI data modeling (Machine learning) and automatic annotation. Bellow image explains the workflow architecture.



Now the crucial part comes into the picture. With the data we are developing semantic web applications with data visualization to view how text and objects are related together. The visualization part look like below picture (From: our Semantic Blumen bach project).



There are many ways to model relationships between ancient texts and objects. Creating statements with RDF triples is one of them. In the workshop ‘Exploration of Text-Object relationship with Semantic Web'. The continuation of this workshop will allow the specific interdisciplinary audience to have a unique forum for discussing, exchanging ideas and experiences while working with different kinds of ancient manuscripts from different domains.

We have endeavoured to discover and render visible the innate connections of any ancient documents (History, Literature, Social science related documents)with the physical objects. In this workshop i am giving a hands on training on a careful evaluation of existing frameworks for describing objects and texts, we decided to test the Scientific Communication Infrastructure (WissKI) for this purpose (Doerr M. Stead S. 2012). During the workshop, new modules have been developed, and a workflow to connect extracted knowledge from different kinds of ancient documents (maybe language independent is also possible) texts with metadata of objects has been established. The data modelling, the ingest workflow, and project evaluation in the context of the ongoing discussion about Linking TEI and CIDOC Conceptual Reference Model are subjects of this article.

In addition, directions like for example the incorporation of semantic technologies for ancient healthcare manuscripts need to be explored with the Object and entity relationship establishment. Decision support, and semantically enhanced AI approaches should be further explored in the workshop. I hope this workshop will offer a fruitful environment for these ideas to mature leading to the ultimate goal of improving the results from the other scientific Arabic documents.

We have endeavoured to discover and render visible the innate connections of any ancient documents (History, Literature, Social science related documents)with the physical objects. In this workshop i am giving a hands on training on a careful evaluation of existing frameworks for describing objects and texts, we decided to test the Scientific Communication Infrastructure (WissKI) for this purpose. During the workshop, new modules have been developed, and a workflow to connect extracted knowledge from different kinds of ancient documents (maybe language independent is also possible) texts with metadata of objects has been established. The data modelling, the ingest workflow, and project evaluation in the context of the ongoing discussion about Linking TEI and CIDOC (Eide Ø. Ore C. 2006) conceptual Reference Model are subjects of this article.

Reference:

Ciula A. Eide Ø. (2014). Reflections on cultural heritage and digital humanities: modelling in practice and theory. DATeCH ‘14 Proceedings of the First International Conference on Digital Access to Textual Cultural Heritage. ACM International Conference Proceeding Series, pp. 35–41. doi:10.1145/2595188.2595207.

Doerr M. Stead S. (2012). Harmonized models for the Digital World: CIDOC CRM, FRBROO, CRMDig, EDM. Slides presented at the CIDOC CRM SIG meeting in CIDOC2012, Helsinki, Finland, June 10. http://www.cidoc-crm.org/docs/CRM_for_TPDL.ppt

Eide Ø. (2013). Ontologies, data modelling, and TEI . In Ciotti F. Ciula A. (eds), The Linked TEI: Text Encoding in the Web. Abstracts of the TEI Conference and Members Meeting 2013. Rome: DIGILAB Sapienza University & TEI Consortium, pp. 26–30.

Eide Ø. Ore C. (2006). TEI, CIDOC-CRM and a possible interface between the two. Digital Humanities 2006. The First ADHO International Conference. Universiteé Paris-Sorbonne. http://www.allc-ach2006.colloques.paris-sorbonne.fr/DHs.pdf (accessed September 2014)

.