



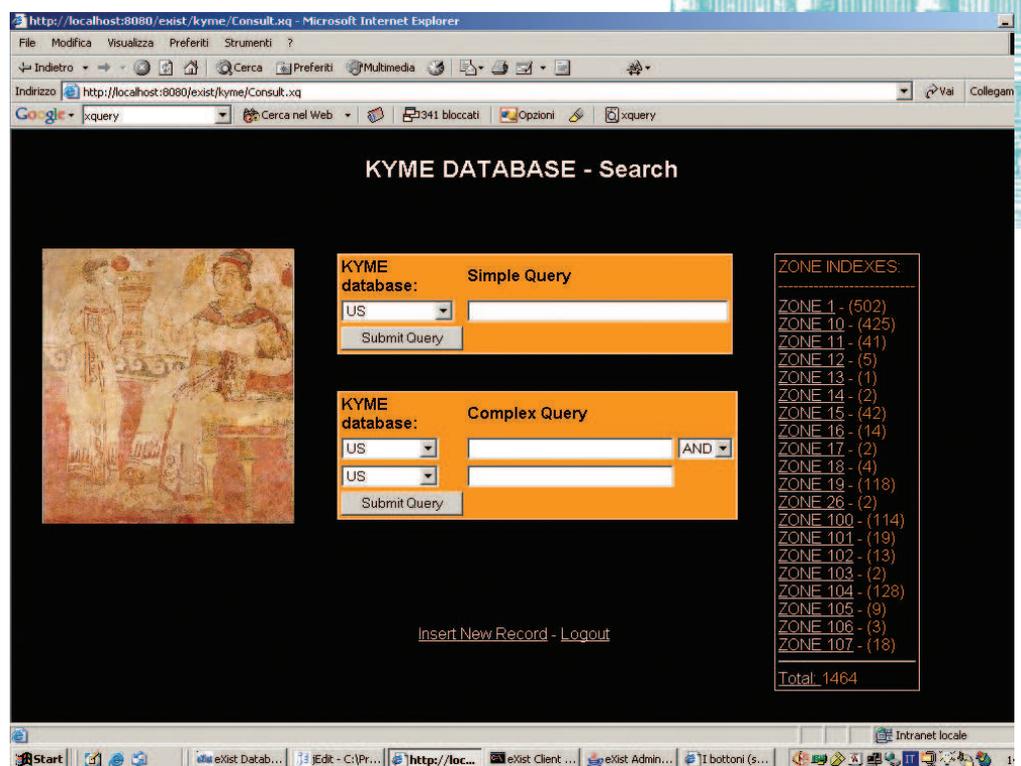
# ARCHAEOLOGICAL DOCUMENTATION FOR THE SEMANTIC WEB

**D**ifferent data sources Archaeological documentation is definitely going digital, but this trend may not be able to solve the problems arising when it is desired to perform a cross-archive search. What is in theory made possible by the support of IT, namely the possibility of managing effectively huge and diverse data archives, is often frustrated by the different structure such archives were given by their creators. This showcase aims at showing that such integration is in fact possible, with an already available technology which substantially improves the way digital archaeological data have been handled as yet. The showcase also will consider existing paper documentation and will show how it can be integrated with digital archives.

**A**pplications The tools used to develop the showcase are freely downloadable on the Internet and are based on public domain standards. It is rather easy to convert existing databases as shown by the application to the excavation database of Cumae, a Classical site in Southern Italy and to a set of excavation databases and excavation diaries from a Medieval site. Digitization and encoding of nineteenth and early twentieth century reports of archaeological collections and sporadic finds in Florence has been successfully tested as well, while the encoding of data on collections of Norwegian archaeological museums, is being mapped to this new system. The case studies total more than 10.000 records.

The management of images and/or drawings is still under development; in some cases it has been tested using SVG, an XML web-compliant vector format for drawings. As far as excavation data are concerned, a system is being developed at Kent University to produce automatically the Harris matrix from the stratigraphic information encoded into the records.

**S**tandards The showcase aims at conjugating standardization – that is the compliancy with the CIDOC-CRM, the standard for museum collections developed by the ICOM documentation committee and already



The search page for the Cumae archive

accepted as an ISO draft – with easiness of use and flexibility. However, it is planned to ‘tokenize’ the proposed encoding in order to guarantee cross-archive interoperability leaving to researchers the choice of record details, tailoring them to the specific researcher’s needs. Multi-lingual issues are also addressed: when existing archives are in different languages, the system enables cross-searches through standardized description of the archive content.

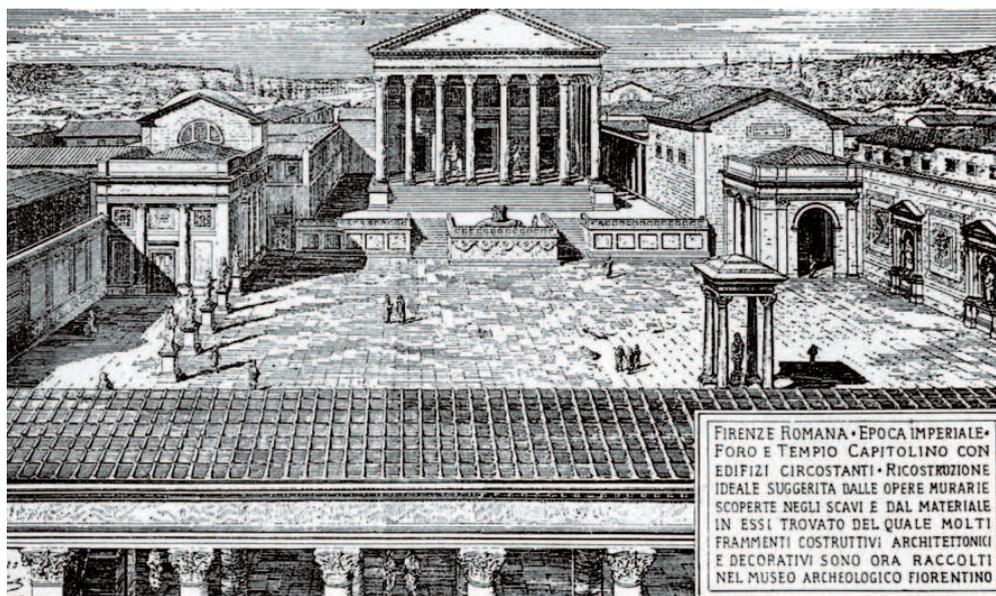
**T**he Semantic Web This buzz-word is used here to indicate the perspective assumed by the proposed methodology: archives may be accessed through the web by means of a browser and they are created in such a way that an intelligent web agent may search them and find relevant information according to user specified criteria. As yet, the system allows cross-archive searches so that, for instance, data concerning ceramics may be searched in dif-

ferent excavation archives (e.g. referring to different campaigns), museum collections and old antiquarian reports. Such searches may take place remotely over the Internet on distributed archives.

**Technical details**

Archives are encoded using an XML CIDOC-CRM compliant structure. Existing databases may be easily converted with no loss of information. The search engine is based on eXist, an Open Source native XML DBMS. Data are organized in collections (corresponding to individual archives) with a hierarchical structure, and each collection may be searched separately or at any chosen aggregation level: e.g. all collections pertaining to a site may be grouped together in a super-collection while archives maintain their individuality, with a directory-like structure.

**Partners** The system is being developed by several partners at different locations to test it under as many diverse conditions as possible. Other contributors are



Roman Florence in the reconstruction of Corinto Corinti (early XX century), one of the old archaeological reports considered as a case study

in the meanwhile proposing additional case studies. The core partners are:

- ▶ PIN, Italy
- ▶ University of Oslo, Museum Project, Norway
- ▶ University of Kent, UK
- ▶ Paveprime, UK
- ▶ University of Naples "L'Orientale" – CISA, Italy



**Interested?**

Are you interested in this showcase? Do you think that this approach can help you in creating effective Cultural Heritage presentation projects or can be integrated in new research projects? Please contact Prof. Franco Niccolucci ([niccolucci@unifi.it](mailto:niccolucci@unifi.it)) of PIN at +39 0574 602513.

EPOCH is a Network of Excellence on Intelligent Cultural Heritage within the IST (Information Society Technologies) section of the Sixth Framework Programme of the European Commission. EPOCH showcases demonstrate innovative solutions and technological integration for target application areas in the Cultural Heritage domain. As they are created with real world content, they stimulate creative thinking about the use of the technologies in Cultural Heritage, and are used to validate new technological approaches with key stakeholders in the Cultural Heritage domain. For more details, visit the project web site:

[www.epoch-net.org](http://www.epoch-net.org)

EPOCH is funded by the European Commission under the Community's Sixth Framework Programme, contract no. 507382. However, this leaflet reflects only the authors' views and the European Community is not liable for any use that may be made of the information contained herein.

