eCHASE: Exploiting Cultural Heritage using the Semantic Web

electronic Cultural Heritage made Accessible for Sustainable Exploitation

Europe's digital cultural heritage content has tremendous exploitation potential in applications such as Education, Publishing, e-Commerce, Public Access and Tourism. Value is hugely amplified if the content can be aggregated, repurposed and distributed at a European level.

The eCHASE project is using semantic web technologies to enable sustainable business models based on access and exploitation of digital cultural heritage content.

P. Sinclair, P. Lewis and K. Martinez
Electronics and Computer Science,
University of Southampton,
SO17 1BJ,
United Kingdom
pass,phl,km@ecs.soton.ac.uk

M. Addis, A. Pillinger and D. Prideaux
IT Innovation Centre,
Southampton,
SO16 7NP,
United Kingdom
mja,agp,djp@it-innovation.soton.ac.uk

Further information at:
http://www.echase.org

Content product producers (publishers, education product providers) load the exported material into content authoring packages where high quality, editorial product can be developed. This could include books, DVDs or web sites.

For example, an e-Learning module developer creates a virtual exhibition of renaissance art, which itself contains links back to the original sources of information as well as a structured series of learning tasks to accompany the book and DVD.

Content product producers provide feedback to content holders about what content they want to access and use.

Once a collection has been completed it can be exported as a complete package for further use in the content creation process. These bundles will contain the media and associated metadata in XML format.

Material is sourced from cultural heritage and content holding organisations across Europe, including museums, libraries and archives.

The metadata is provided in various formats, ranging from database dumps and XML to Microsoft Excel spreadsheets and CSV files. Metadata can also be provided embedded in the media itself. Many media formats are accepted, including images, video and sound files.

CIDOC CRM

The CIDOC Conceptual Reference Model (CIDOC CRM) is a reference model for the interchange of information in the cultural heritage domain.

The eCHASE project is using the CIDOC CRM as the common metadata schema to cover the different metadata repositories from our partners' collections.
http://cidoc.ics.forth.gr/

SKOS

SKOS provides a model for expressing the basic structure and content of concept schemes, such as thesauri.
http://www.w3.org/2004/02/skos/

Legacy Data Selection

Content Product Creation

Semantic Mapping

Export

Data Transformation

Taverna

We are employing the Taverna Workbench, a service oriented workflow system, to perform the mapping from different metadata schemas.
http://taverna.sourceforge.net/

Content Selection and Annotation

Knowledge Base Creation

Having transformed the legacy data into a consistent, well-structured schema, it is converted into RDF. The RDF model is used by the mSpace browser component of the webapp.

Search, Retrieval and Navigation

User content is managed by a personal light box tool, which gives authors and experts a mechanism to collect, organize and annotate information for later use.

Items found through the search engine with the web interface can be added to one of the users light box collections to either be used for future searches or to be exported as an item in the content package the user is building.

Textual queries can be run on the data in the unified metadata through a form based search.

A media engine service hosts the media and provides content-based querying facilities, including searches based on colour or texture.

Textual and content based queries can be combined: for example one can search for red chairs.

SRW - Search/Retrieve Web Service

The framework supports a search and retrieval protocol based on the SRW specification developed by the z39.50 community, providing a "search" operation to handle common query language (CQL) queries. The SRW supports queries based on each collection's legacy metadata schema and on the unified database schema based on the CRM.

Thesauri navigation is provided by presenting a thesaurus tree structure to the user, such as concept hierarchies or gazetteer place names.

Gazetteer latitude and longitude data is also used for visualizing creation place information for each object. The map visualization interface is provided by the Google Maps API.

Browsing is provided by an mSpace interface, an interaction model designed to allow the navigation of multi-dimensional spaces.
http://www.mspace.fm