Trends in Digital Cultural Heritage Management and Applications

Stavros Christodoulakis

MUSIC/TUC Laboratory, School of Electronic and Computer Engineering

Technical University of Crete

Stavros@ced.tuc.gr

Abstract

We present some recent trends in the field of digital cultural heritage management and applications including digital cultural data curation, interoperability, open linked data publishing, crowd sourcing, visualization, platforms for digital cultural heritage, and applications. We present some examples from research and development projects of MUSIC/TUC in those areas.

Digital Cultural Data Curation

A curator is responsible for the acquisition and care of cultural objects. Typically he makes decisions regarding what objects to collect, he oversees their care and documentation, he conducts research based on the collection, he arranges proper transport of cultural exhibits, he communicates with the public and the research community through exhibitions and publications, etc. Curation of digital cultural heritage data involves among others making decisions regarding what digital cultural heritage data to collect, oversees cultural heritage data care and documentation (metadata), conducts research based on the collection, plans for proper packaging of the cultural heritage data for reuse, takes actions for the long term reusability and value of the cultural data, and arranges for sharing the cultural data with the public and researchers.

Cultural Data Curation is done for humans, users with interests in culture, cultural researchers, learners, or even commercial users that aim to provide value added products or services using the cultural data. Systems that help the data curators to provide such functionality should provide facilities for collection management, collection and object management, web browsing and viewing, and search functionality. Proper metadata management is essential to effectively support these functionalities.

Interoperability

A very important consideration of the cultural heritage institutions today, especially for the medium and small size institutions is discoverability in the web. To increase the knowledge and the appreciation of the value of the cultural heritage institution its better collections or exhibits, at least, should be discoverable by search engines and domain specific engines or have some presence in large cultural portals. Such presence in portals or discoverability by domain search engines typically implies functionality for interoperability [18]. An important form of data interoperability is metadata mapping between the metadata of the cultural institution and the metadata of the portal or the search engine. In other cases more complicated wrappers have to be built to provide retrieval language, schema, and data mappings.

Semantics, Open Linked Data Publishing, Data Provenance

Cultural data curation is strongly related to cultural data semantics. The data semantics and the data interrelationships should facilitate humans for understanding the cultural objects. Adding value to the cultural data typically requires following a systematic approach to modeling and metadata creation, publishing the data as Linked Open Data, and linking them to other important data sources. The publication of the cultural data as Linked Open Data enables the creation of user communities of domain experts. RDF is used as a target model for publishing the Open Linked Data.

The process of Open Linked Data publishing typically involves finding the appropriate resources to be linked, discovering their Schema, proving rules for transformation of the cultural institution data to the Linked Open Data format, providing rules for linking with the external sources, storing the transformed and linked data in an internal repository, and publishing them in the web. A challenge in this process is finding appropriate models for automating as much as possible the linking processes. Functionalities for integrating prior knowledge, exceptions, and human intervention are also very important ([6],[7]).

Data provenance for cultural data curation is concerned with the explanation of results in questions like why was a particular link produced, why a link is of a specific type, what rules were used to derive a link, who created the link, how trustworthy the link is. It may also be concerned with data mining aspects of the use of the cultural data in order to better satisfy the user requirements. System functionalities should support modeling and querying on provenance data.

The Natural Europe infrastructure

In the context of the Natural Europe project we have developed methodologies, software architectures and systems to support Natural History Museums for their web presence, their interoperability with major international metadata providers and search engines and for publishing their data as Linked Data ([12], [13]). The systems have been installed and used in six important European Natural History museums. Each museum node is provided with a multimedia authoring tool (MMAT), a Cultural Heritage Object (CHO) repository and with a Vocabulary Server facilitating the complete metadata management lifecycle ingestion, maintenance, curation and dissemination of CHOs. The infrastructure also supports the migration of legacy metadata into the node. The application profile of CHOs has been created through an iterative process with the museum domain experts and it is a superset of the Europeana Semantic Elements (ESE) metadata format, thus providing a direct interoperability with the central European Digital library (Europeana). The CHO repository manages both content and metadata and adopts the OAIS Reference Model ([8])for ingestion, maintenance and dissemination of information packages. The Vocabulary Server supports any taxonomic classification that the museum may use. The ingested taxonomies follow the SKOS format which is a leading international standard based on the Semantic Web principles for representations of Thesauri, Taxonomies and other types of controlled vocabularies. The controlled vocabularies provide strong support for the curation, indexing, retrieval, autocomplete functionality, etc. For Natural History an important vocabulary is is the Catalogue of Life (CoL) which contains 1.4 millions of species and their relationships. We have expressed the taxonomy of CoL to SKOS using the CoL annual checklist and a D2R server.

An Access Module provides a number of services that allow selective harvesting of metadata from external entities through an OAI-PMH interface. The museums can be seen individually or through a federation. The Access Module is used to harvest the metadata to the federal node, to Europeana, as well as to establishing connections with major biodiversity networks such as GBIF and BIOCASE [5]. The BIOCASE network is based on a very involved Schema (ABCD Schema) which describes nearly 1200 different concepts. We developed in cooperation with museum experts mappings between the ABCD schema concepts, and wrappers to be used by BIOCASE to access the XML databases of natural Europe (the BIOCASE wrappers assume relational dbms underneath). The wrappers developed follow a layered architecture so that they can be easily adapted for other XML data sources. To support the Semantic Web presence of each museum individually we have described in OWL the CHO Application profile of Natural Europe. The resulting Natural Europe ontology references well known ontologies/schemas (like DC, FOAF, Geonames, SKOS) and has been aligned with the Europeana Data model (EDM [4]) supporting interoperability with the Europeana Semantic Layer. The publication process [13] involves establishing links to the external RDF data sets, conversion of the XML data to RDF, maintainance, publishing and dissemination of RDF data. The Semantic Infrastructure allows highly expressive queries combining knowledge from distributed resources like 'find photos of endangered species of genus "Bufo" in neighboring countries of Greece' which combines information from Natural Europe, DBpedia, CoL/Uniprot, and Geonames.

Crowdsourcing

While cultural institutions have specific processes and scientific personnel for the curation of data, in some scientific domains it is simply too expensive to gather all the data needed. A typical field is biodiversity. There is not enough personnel and budget in natural history museums and similar organizations or research centers to gather the required data. Crowd sourcing may offer a solution to the problem. Appropriate social procedures should be deployed to distil the quality of data gathered with crowd sourcing processes. We are implementing crowd sourcing infrastructures that facilitate community interaction for knowledge distillation in the context of nature observations ([17], [11]).

Visualization and Interaction

Good visualization and user experience is important for attracting and keeping visitors to the exhibits of a cultural site. Web applications today can integrate in a standard manner multimedia data types (pictures, video, animations, sound) and allow their presentation by any web capable device. Integrations of diverse mobile devices such as tablets and mobile phones becomes very important for many cultural applications like tourism.

Humans get a great satisfaction with the contact and interaction with other humans. Interfaces that enable human interaction and joined activities greatly enhance the user experience of the applications.

Contextual Viewing and Understanding

The integration of diverse sensors with the mobile devices is particularly important. Sensors today may be used to provide much contextual information like the GPS location, compass direction, direction of device movement, RFID, and many other sensor types become available for integration. These sensors enable the association of context with the user, such as his location on a map or a diagram together with the surrounding objects, including the cultural objects. Social context (who is near you), is also possible and can be used for enhancing the satisfaction of a cultural experience, for scheduling, for security, etc.

The mobile devices, the sensors that are integrated on them, and the access to the internet and the semantic web enable the mobile device to become a window on the world. We have designed and implemented a framework and a system that allow identification and visualization of the semantics of distant objects in nature, which exist in pictures taken by mobile devices [3]. Semantic objects in nature are represented on 3D maps and tagged with semantic and spatial descriptors. The system functionality creates, on top of pictures, a layer that allows semantic user browsing in augmented geospatial images, semantic object identification in pictures, object interaction, contextual map visualization, and interlinking of information with digital libraries and the web.

Of significant importance are the recent efforts of providing contextual information for a user walking within a building. Sensors and software are being developed to allow a person walking in any floor of a building (where GPS may not be available) to have a contextual knowledge of his environment. Apple is launching a vast project to map the inside of every large building it can [1].

4D visualization, Augmented Reality

3D visualizations of cultural objects facilitate the users to understand how an object (like a fort that does not exist any more) looked like. Digital reconstruction of buildings or even cities may be a difficult job that may have to reconcile conflicting views by different teams. Contextual support in the mobile devices allows powerful contextual views of (parts of) the reconstructed objects from the location and point of view of the user.

Reconstructions may be enhanced with augmented reality aspects, like the semantic multimedia annotation of various parts of the reconstruction which provide explanations of the functionality or the way of use of the various objects in the reconstruction.

4D visualizations try to integrate the time parameter in such visualizations. For example, some European cities have a history of thousands of years, and they went through periods of disasters which were followed by periods of new expansion. Understanding the evolution of the city through the millennia is important for cultural heritage.

Another example where 4D visualizations are useful for digital cultural heritage is the reconstruction of historical events that took place in space and time. The existing geovisualization tools (like Google Earth) may be used to build on top of them the visualization in time. In [16] we describe ENVISUGE, a model and a system that allows the visualization of time dependent geospatial events on top of Google Earth. As an example the battle of Marathon between Greeks and Persians (490 BC) has been described in the model and can be visualized on top of Google Earth.

Portals, Federations, Platforms

Different types of software infrastructures can be used for exploiting the cultural data in the web.

Portals and Federations

Cultural Heritage Portals typically integrate cultural heritage metadata from one or more cultural heritage institutions using a standard metadata model, and they allow browsing, searching and metadata viewing for the cultural objects. Europeana is currently a typical example of a portal. It aims to be a central digital repository for cultural heritage metadata in Europe. Europeana uses a standards set of meta data for integration with the cultural institutions. The cultural institution has to map (a subset of) its own metadata for each cultural object to the metadata types used by Europeana. Europeana provides tools that facilitate this process. The mapped metadata are checked by Europeana for conformance and are stored for answering user queries and browsing.

Federations allow access to independently developed cultural sites, typically on a certain application domain. Queries are submitted to a federation Schema and results from the federation sites are returned. The nodes of the federation may keep their own data, in which case a mapping of the federation Schema to the local Schemas allows the user queries to be evaluated on the node data bases. For all relational federation nodes much research has been done in this area. For other models (OWL, XML) or mixed model environments more research is needed ([2], [9]).

Cultural Platforms

While portals emphasize the integration of data from various sources, platforms emphasize services offered for building applications by the user community. The services offered can be in the form of SOAP or lighter REST services. A cultural platform (in addition to possibly having the capabilities for integrating cultural data from various sources), will facilitate the building of new value adding services to exploit in various ways the cultural data. In general, the services of the platform and the new user added services that operate on the platform can be used again for building new services that are offered to the users of the platform. This is the basis of Web of Web or web 2.0 infrastructure. This infrastructure facilitates the inexpensive, risk sharing applications for the long tail market segments which were not traditionally exploited due to their small size and cost of development.

The new Europeana business plan targets Europeana to become a platform that will support user generated services and will integrate better the Europeana user community, end users, cultural institutions, partners, developers, customers. It will provide shared technical infrastructure for all partners to use. A service oriented architecture will form the basis of the platform and it will allow applications B2C (Business to Customer), B2B (Business to Business) and B2N (Business to Network) to be developed. In Europeana Labs the developers will find data, content, technology, documentation that they use to build new apps, applications and services. They will also find apps developed with Europeana data, and support tools for collaborative work, use of the Europeana API and Linked data. Europeana as a platform will have to provide mechanisms for profit sharing with the cultural institutions.

Applications for Sustainability

Infrastructure Benefits and Sustainability of Digital Cultural Heritage Institutions

The digital heritage infrastructure provides multiple benefits to the community. For the galleries, libraries, museums, archives there are benefits from the improvement of the communication of the institution to the public and the professional users. An additional benefit is that it reduces the cost of finding the institution offerings for those seeking the development of applications and services based on culture. For the general public that is interested in culture it provides a remote visualization of important institution exhibits, and thematic exhibits. For the tourism sector it can provide applications and guide books which are particularly important for the less known regions. For the creative industry it facilitates the publishing on arts, culture, travelling guides, and others. It facilitates journalists looking for historical info, as well as artists and designers and game developers. For the educational institutions, teachers, scholars, students, it provides accessible digital material at all times from everywhere.

Some of the most important economic benefits of the digital cultural heritage infrastructures according to Europeana Business Plan are the cost savings for institutions and the additional income from tourism, while at the same time they contribute to the increase of the community cohesion and raising awareness and knowledge. The benefits to specific communities like learners and scholars is also very high. Workflow support for the repeated contacts of those communities with the institutions are of great benefit to both sides, and supports the sustainability of the cultural institutions.

Experiencial Visits

Modern technologies enable innovative applications for institution visits like museum visits, open space visits, and smart city visits. Many applications are enabled with the rapid expansion of internet of things. Devices that detect the human presence and interact with the user and the user device allow much better and emotional user experiences. Imagine the reaction of a child if in a natural museum visit a species recognizes the near presence of the child and it produces the sound that it makes in real life. Imagine also facilities that simulate and animate the behavior of a hunting species in the presence of a human.

The mobile devices with GPS, direction and other sensor capabilities also allow contextual personalized storytelling in museums and open spaces. The storytelling can adapt to the user interests, knowledge, path followed, past behavior, presentation device and preferences, in the presentation of surrounding actual or reconstructed cultural heritage objects or events happened. Such applications are important for tourism.

Thematic crowd-sourcing applications in specific cultural domains like nature are also enabled by the mobile devices and sensors, and they are also important for tourism ([11], [17]).

Learning

Learning is one of the most important applications of digital cultural heritage. Infrastructures for effectively planning domain learning and visits to cultural heritage institutions can support pre-visit, during-visit and post-visit experiences to cultural institutions or spaces with a variety of tools for many different learning paradigms. Appropriate metadata management is important to support these processes. Teacher and student community participation and evaluation to learning activities are also important for the continuous renewal and improvement of the learning methodologies.

Octopus [10] is a tool that supports learning planning, association of the plan with available resources (such as cultural data) and collaboration between instructional designers and teachers; editing and sharing of educational practices (templates) and educational scenarios including reuse of their parts through drag and drop; interoperability with external repositories and environments exploiting available designs but also sharing educational practices and scenarios developed within the system using their APIs and implementing protocols such as OAI-PMH and Open Search. Octopus has been used in large projects (like ODS) supporting many diverse learning environments and large user communities from teachers to children to University students.

Games for learning purposes are an established approach for learning. Games in nature are facilitated by the modern mobile devices and Sensors. Real time application support for learning purposes, (in general for the design and enactment of collaborative learning scenarios) as in [14]is important for learning.

Cultural games are also significant for intergenerational learning and for avoiding exclusion through remote collaboration. Shadow Theater is popular in many countries around the world. eShadow [15] is a web based storytelling tool that can be used from both adults and children in order to create, record, share and watch digital shadow theater plays. Children in Greece watch traditional shadow theater plays, learn about Shadow Theater in school and play with shadow theater puppets. Because of this real life experience new possibilities for digital culture emerge such as the enactment of intra-family communication scenarios that promote intergenerational bonding and playful learning. These new opportunities for intergeneration bonding that overcomes the physical separation of children and their grandparents is important for children's development and contributes to the well being of the elderly as well. Karagiozis and Akritans shadow theater creations based on traditional art and learner creations, accompanied with creative music performances proved to be very popular to children.

Conclusions

We have outlined some trends in the management and the applications of the digital cultural heritage. Technology changes like the profitability of mobile devices and sensors open up significant advances in the application experience for digital cultural heritage applications in creativity, tourism, learning, and others. The software infrastructure for effectively supporting these applications is significant and involves many curation tasks and workflows, as well as significant visualization and interaction and platform capabilities. We have overviewed some important functionalities needed, the benefits of this infrastructures for the cultural institutions and the society, and we have outlined some modern applications of digital cultural heritage.

References

1.Apple is launching a vast project to map the inside of every large building it can: http://www.businessinsider.com/apple-indoor-mapping-project-and-ibeacon-2014-6

2.Bikakis N., Tsinaraki, C., Stavrakantonakis, G., Gioldasis, N., Christodoulakis, S.: TheSPARQL2XQuery Interoperability Framework, WWW Journal 2014.

3.Christodoulakis, S., Foukarakis, M., Tsinaraki, C., Ragia,L., Kanellidi, E.,: Contextual geospatial Picture Understanding, management and Visualization, the proceedings of the 11th International Conference on Advances in Mobile Computing and Multimedia, MoMM 2013.

4.Europeana Data Model Definition V.5.2.3. http://pro.europeana.eu/documents/900548/bb6b51df-ad11-4a78-8d8a-44cc41810f22.

5.Global Biodiversity Information Facility. http:// www.gbif.org/

6.Hassanzadeh, O., Pu, K., Yeganeh, S., Miller, R., Popa, L., Hernandez, M., Ho, H., : Discovering Linkage Points over Web Data, Proceedings VLDB 2013.

7.Hendler, J., Ding, Y.: Publishing and Using Cultural Heritage Linked Data on the Semantic Web. Morgan& Claypool Publishers series (2012)

8.ISO 14721:2003 Open Archival Information System (OAIS) Reference Model. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=24683 (2003)

9.Makris, K., Bikakis, N., Gioldasis, N., Christodoulakis, S.: SPARQL-RW: transparent query access over mapped RDF data sources. In: Proceedings of the 15th International Conference on Extending Database Tech- nology, EDBT. ACM dl (2012).

10.Mylonakis, M., Arapi, P., Moumoutzis, N., Christodoulakis, S., Ampatzaki, M.,: Octopus: A Collaborative Environment Supporting the Development of Effective Instructional Design, Proceedings International Conference on E-Learning and E-Technologies in Education, ICEEE2013.

11.Skevakis, Tsinaraki, Trochatou, Christodoulakis: "A Crowdsourcing Framework for the Management of Mobile Multimedia Nature Observations" IJPC, 2014 (to appear).

12.G. Skevakis K. Makris V. Kalokyri P. Arapi, S. Christodoulakis: Metadata Management, Interoperability and Linked data Publusing Support for natural History Museums, IJDL, 2014.

13.Skevakis, G., Makris, K., Arapi, P., Christodoulakis, S.: Elevating Natural History Museums’ Cultural Collections to the Linked Data Cloud. In: Proceedings of the 3rd International Workshop on Semantic Digital Archives (2013).

14.Stylianakis G., Moumoutzis, N., Arapi, P., Christodoulakis, CoLearn: Real Time Collaborative Learning Environment The Second International Conference on E-Learning and E-Technologies in Education (ICEEE2013), Sept. 23-25 Poland 2013.

15.Christoulakis M., Pitsiladis A., Moraiti A., Moumoutzis N., Christodoulakis S. : EShadow: A Tool for Digital Storytelling Based on Traditional Greek Shadow Theater, proceedings Intenational Workshop on Intelligent Digital Games for Empowerment and Inclusion, Foundations of Digital Games Conference, 2013.

16.Tarantilis, N., kazasis, F., Gioldasis, N., Christodoulakis, S.: Event Visualization on Google Earth, Proceedings SMAP 2011.

17.Tsinaraki, Skevakis, Trochatou, Christodoulakis: "MoM-NOCS: Management of Mobile Multimedia Nature Observations using Crowd Sourcing" Proceedings MoMM2013.

18. Tsinaraki, C., Polydoros, P., Christodoulakis, S. : Integration of OWL Ontologiew in MPEG-7 and TV-Anytime Compliant Semantic Indexing, Springer LNCS, V 3084, 2004, pp. 398-413.