IST-2002- 507382

EPOCH

Excellence in Processing Open Cultural Heritage

Network of Excellence

Information Society Technologies

D3.3.3: 3rd 6-monthly EPOCH Pipeline Description

Due date of deliverable: 15 September 2005
Actual submission date: 15 October 2005

Start date of project: 15 March 2004
Duration: 4 Years

University of Leuven

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	X

1 Table of Contents

2 Work package 3, activity 3.3: Common infrastructure

This deliverable reports on the activities regarding the common infrastructure in the third semester of the project. After the first year of the project, an update to the work plan was prepared, taking into account the results and experience gained. This JPA30 proposal was discussed at the first review meeting, but no final decision about it was obtained from the EC until now. This delay and the uncertainty about the continuation of the project have interfered with the work because quite some changes were proposed for WP3 in this semester.

In this deliverable we report according to the original JPA18 planning but at the same time the difference with JPA30 will be mentioned. This accounts for the fact that some of the milestones were not reached according to the original JPA18 timescale.

Apart from the results described below, a number of partners have formed a productive research partnership under the aegis of EPOCH, which in turn has led to further collaboration in project proposals. The results of some of these projects will contribute to the toolbox provided by WP3.

Another significant result of EPOCH collaboration between ARCES (Università di Bologna) and the University of Kent is the recent successful organisation of a workshop entitled *Smart Environments and their Applications to Cultural Heritage*, held at the UbiComp conference in Tokyo, Japan on 11th September 2005. UbiComp is the primary international conference of the Ubiquitous, Pervasive and Ambient Intelligence research communities. The workshop attracted contributors from Europe, USA, Japan, Australia, Korea and Taiwan, and the papers have been published as an EPOCH publication.

2.1 Objectives for the third semester

WP3.3 aims at establishing and maintaining a Common Infrastructure which will form a substrate upon which to build particular end-user applications and showcases. This activity defines the architecture, components and design guidelines for this common infrastructure.

In the proposed JPA30, this was further clarified. The goal is the creation of a toolbox of inter-operable modules for producing applications involving digital versions of tangible cultural heritage. The toolbox will be fed from four sources: (1) pre-existing tools and tools developed outside the network, (2) developments from within WP3.3 in order to provide a consistent and cohesive framework, (3) tools polished within the scope of the Showcase activities, and (4) NEWTON projects to fill in missing components.

The proposed milestones for this semester, as described in JPA18, are the *technical benchmarking of tools* (M3.3.4) and the release of the *first set of open software tools* (D3.3.4). However in the proposed JPA30, the milestone about technical benchmarking has been dropped and deliverable D3.3.4 has moved to month 30.

2.2 Overview of the work

In the first year, the concept of the common infrastructure was defined and a list of available tools for cultural heritage was compiled (cf. D3.3.2). Central to this infrastructure is the information store. Its usability is determined by fixing (or at least limiting) the interchange formats used.

This can of course not be done for all possible cultural heritage applications, so a set of topics was selected for development under the Common Infrastructure based on the priorities on the first version of the Research Agenda and the subsequent poll among partners about the best strategy to achieve them:

1. tools for conversion of existing catalogues (JER subarea 1)
2. multilingual systems for collection interrogation (JER subarea 1)
3. link repository for 3D primary data (JER subarea 2)
4. multi-modal data retrieval (JER subarea 2)
5. integrity management of presentation data (JER subarea 2)
6. internet portals based on distributed systems (JER subarea 2)
7. contextual cultural information (JER subarea 3)
8. supportive measures for mobile applications (JER subarea 3)
9. 3D file format compendium (JER subarea 4)
10. large cultural and natural heritage sites (JER subarea 4)
11. unified framework for 3D applications (JER subarea 5)
12. avatar standards for cultural heritage (JER subarea 7)

Based on the priorities for the Research Agenda on this list, a distribution of work among the partners was prepared. This distribution of work is described in detail in the revised JPA30 proposal. The partners committed themselves to the following WP3.3 activities:

- University of Brighton [partner 1]: (1) data query tools, (2) a framework for 3D applications; and (3) acting as WP3.3 – WP4.2 go-between.
- PIN [2]: data and knowledge management, especially on (1) a turnkey installation wizard, (2) the analysis of usability and impact, and (3) image- and 3D-based databases.
- Ename [3]: knowledge management in an archaeological and historical context, with focus on (1) sustainability and updating representations, (2) information integration and updating, and (3) linking interpretation and presentation.
- University of Leuven [4]: a web-based tool to upload images and get 3D models from a server installed at the Leuven site, consisting of (1) An upload tool, (2) the server software, and (3) building a 3D model repository.
- University of Braunschweig [19]: 3D modelling and visualisation, in particular (1) the establishment of OpenSG & GML as standards, (2) new house representations, and (3) a new city modelling application.
- IGD [20] and TII [46]: monitor developments in the VR/AR, and (multi-modal) interfacing and human-machine interaction domains.

- University of Bologna [30]: in the mobile/wearable/ambient area: (1) quality vs. power consumption, (2) head-up and head-worn displays, and (3) context management frameworks.
- CNR-ISTI [33]: tools for the processing of 3D data: (1) 3D data editing suite and (2) visualisation of huge data sets.
- ETH [48]: create a grammar-based framework to automatically generate buildings of a prescribed architectural style and tallying with prescribed constraints: (1) sources of possible constraints, (2) the shape grammar rules, and (3) completing the mode with tools for streets and vegetation.
- University of Geneva [49] and EPFL [50]: aspects related to avatars and virtual human: (1) tool inventory completion, (2) formats and standards, and (3) technical benchmarking.
- University of Kent [52]: creating a flexible excavation recording tool and other aspects of the mobile/ wearable/ambient area.
- University of East Anglia [56]: (1) tools for Avatar creation, (2) rendering and animation, and (3) interaction with the user.

These activities will lead to tools and tool chains within the EPOCH common infrastructure. The work on it has started and some tools and tool chains are defined. The ones which are available (or will be in the near future) are described in section 3 and 4.

To conclude the overview of the work, the status of the two JPA18 milestones of this period is described.

2.2.1 Technical benchmarking of tools

As already mentioned in deliverable D3.3.2, benchmarking tools proved to be impractical for the current set of tools.

First of all, it proved to be difficult to compare tools if their functionalities did not precisely overlap. An example was the planned benchmarking of archaeological recording tools, which would all be tested on the basis of the Sagalassos excavation campaign in the summer of 2005. Yet, upon closer inspection, there were too large differences in the functionalities of the tools to get much further than listing and comparing such functionalities. This has been done, but genuine benchmarking calls for running the same experiments on the same data, which proved difficult. Most tools have a (at least slightly) different audience and/or scope, so it's not fair to compare them. For other tools, like 3D scanning devices, more thorough benchmarking is possible, but recent reports were found where such benchmarking results had already been published (e.g., the SparView reports in this case).

Another issue is liability, mainly in the case of commercial systems. Benchmarking is delicate, unless the companies are offered the possibility to operate the systems with their own experts. Otherwise, inferior results can easily be claimed to be due to improper or suboptimal use. This would increase the price and in the case where we tried to arrange such a comparison even for modest subtasks, as in the case of archaeological recording tools (the intended test was 3D stratigraphy recording and bringing in finds), some companies were very reluctant or insisted to await the availability of an upcoming, enhanced product.

Therefore technical benchmarking is removed from the proposed JPA30 and no activity happened on this topic in this semester.

2.2.2 First set of open software tools

As EPOCH has been explicitly requested by the EC not to develop any technology itself during its first year, the reported tools and components have mainly been the result of work done outside of the network, even for the tools offered by the EPOCH partners. Therefore deliverable D3.3.4 should actually be delivered in month 30, according to the proposed JPA30 update of EPOCH. By that time the EPOCH toolbox will have crystallized into its definitive form and a more extensive list of tools proposed by EPOCH will be available. However to comply with the current JPA18 contract, an interim deliverable D3.3.4 has been prepared with the subtitle “version month 18”. It investigates the possibilities and restrictions of open software in the field of cultural heritage, clarifies EPOCH’s attitude towards open software and gives a preliminary list of proposed EPOCH open software tools.

A summary of the tools promised by EPOCH partners at this moment is found in section 3 of this report.

2.3 Future planning

If the project administration problems gets solved by the next VAST-conference, then WP3.3 will organize a workshop at VAST on the open issues of the common infrastructure. We’ll define the tool chains to be implemented during the next months, complementary to the NEWTON activities. We’ll also make more explicit which interchange formats have to be promoted. These interchange formats (the “common” in “common infrastructure”) are not restricted to well established standards, but also other open interchange formats such as COLLADA[1] will be considered.

We are confident that this will bring us back on the JPA30 track by the end of the second year period.

[1] http://collada.org/

3 EPOCH tools

The deliverable D3.3.4 ("version month 18") investigates the possibilities and restrictions of open software in the field of cultural heritage and it formulates an EPOCH opinion on it. The core part of this deliverable consists of a survey of open source solutions specific to cultural heritage. The survey is complemented by an overview of European research projects dealing with open software and by specific tools provided by EPOCH partners.

The list of EPOCH tools, which is summarized below, is by no means an exhaustive or complete list. NEWTONs are not included. Quite some partners already initiated the development and integration of additional tools within the framework of the common infrastructure (cf. the commitments in section 2.2). However due to the uncertainty about the future of the network, these activities are still in their initial phase so they are not mentioned in this report.

The following tools from EPOCH partners are described in more detail in the current deliverable D3.3.4:

- *Tools for mobile, wearable and ambient technologies for archaeology*:
  - FIELDMAP: a field data collection tool, which runs on off-the-shelf PDAs
  - MOBICOMP: an experimental support infrastructure for storing, managing and sharing contextual information in a distributed environment
  - JNET: a tool for visualising, manipulating and editing diagrams of the stratigraphic sequences, or Harris Matrices, derived from archaeological excavations
- *3D-model web service*: a free service on the Internet for registered users to upload images, which are then automatically turned into 3D models, and sent back to the party uploading the images
- *3D Scanning Tools*:
  - MESHALIGN: registration of multiple range maps
  - MESHMERGE: reconstruction of a single 3D mesh out of a set of registered range maps
  - MESHEDIT: performing simple editing actions on the mesh
  - MESHSIMPLIFY: simplification of the (huge) meshes produced by 3D scanning devices, by removing mesh vertices in a controlled manner
  - WEAVER: management of a set of images and building up a texture map wrapped around the 3D model
- *Combining GML (Generative Modelling Language) and OpenSG*: exposing the OpenSG API to the GML language, and integrating the GML runtime engine into the scene graph engine combines the strengths of two technologies
- *Avatar manipulation*: classes for avatar loading, rendering, and animation in OpenSG and VHD++

Some of the other tools which are under development:

- CIDOC-CRM compliant response generation from natural language queries in English
- Design of multi-modal interfaces, such as a portable dome which provides input for image based modelling

4 EPOCH tool chains

The development of tool chains is essential in the Common Infrastructure activity to test the interoperability of tools. The evaluation of these tool chains will provide the basis for the definition of the data exchange requirements and it provides their validation. If some existing tools prove to be incompatible qua data exchange requirements, we'll focus on developing automatic import/export tools for these tools rather than on developing replacements.

At the same time, these tool chains can be used as little showcases or demonstrators to convince people of the possibilities of tools and their combination in the field of cultural heritage.

Two examples of tool chains, which were developed by partners, are given below. More tool chains were discussed between partners, but they were postponed for a couple of months because of the uncertain future of the project. In many cases partners already started working on their part.

Additional tool chains will also be added by NEWTON projects.

4.1 A CIDOC-CRM compliant XML database for archeological data

XML is a cross-platform, software and hardware independent tool to store or transmit information. Because it uses plain text files to store data, it is suited for long term data preservation and a text editor suffices to create or modify an XML document. Because of these advantages almost all new standards and exchange formats build on it.

For the storage of archeological data a native XML database is preferred. It provides a search engine and other tools for the management of XML documents. The documents themselves preserve their XML nature and structure when stored or retrieved. Several implementations were evaluated and finally the Open Source EXIST[2] database was selected. It is written in Java, is platform independent and has support for XQuery and XPath. It can run as a stand-alone database server, as an embedded Java library or as part of a web application. XML documents are managed in hierarchical collections, similar to storing files in a file system. Collections are not bound to a predefined schema or document type, which allows us to use a CIDOC-CRM compliant XML Schema.

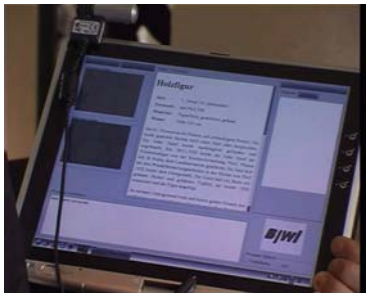
The ontology related work has been done using PROTÉGÉ[3], a free Open Source ontology editor and knowledge-base framework. It is based on Java, is extensible, and provides a foundation for customized knowledge-based applications.

Both tools have been used in a case-study to generate an ontology for archaeological records and discuss results with archaeologists. It was based on the Cumae archive of some 2500 excavation records. All data was converted to clean CIDOC-CRM compliant XML before storing it into the XML database.

[2] http://exist.sourceforge.net/

[3] http://protege.stanford.edu/

4.2 Interactive Museum Guide



An example of an application build from a tool chain is the prototype of an interactive museum guide. It runs on a tablet PC that features a touch screen, a webcam and a Bluetooth receiver. This guide recognizes objects on display in museums based on images of the latter which are taken directly by the visitor. Furthermore, the computer can determine the visitor's location by receiving signals emitted from Bluetooth senders in the museum. This information is used to reduce the search space for the extraction of relevant objects. Hence, the recognition accuracy is increased and the search time reduced. Moreover, this information can be used to indicate the user's current location in the museum. The prototype has been demonstrated to visitors of the Swiss National Museum in Zurich.