EPOCH Conference on Open Digital Cultural Heritage Systems (2008)
David Arnold, Franco Niccolucci, Daniel Pletinckx, Luc Van Gool (Editors)

A Common Infrastructure for Cultural Heritage Applications

L. Van Eycken[1], B. DeKnuydt[2], and L. Van Gool[1]

[1]Center for Processing Speech and Images (ESAT–PSI/VISICS) – IBBT, K.U.Leuven, Belgium
[2]IMEC vzw – IBBT, Belgium

Abstract

This paper explores the infrastructure needs for cultural heritage applications. Small dedicated applications as well as very large projects are considered. A unified approach for managing cultural heritage information is proposed to handle storing and exchanging data. An implementation demonstrates its use in two different cultural heritage applications.

1. Introduction

The jointly executed research of the Epoch project aimed at providing the infrastructure and the tools to build cultural heritage applications from. Starting from stakeholder needs and implementation experiments, we looked for a common infrastructure to support the various applications. An important aspect of cultural heritage applications is their diversification. This relates not only to the kind of data (descriptive text, semantic relations, location information, sound, image and 3D representation, . . .) but also to the way the data is used (preservation, interpretation, visualisation, exhibition, web catalogues, . . .).

Let's consider as an example a large historic site, like the Valley of the Kings in Egypt. The amount of objects that need a description is enormous: there are about 627 tombs, with hundreds of noteworthy items (sarcophagi, masks, wall-paintings . . .) per tomb. Each item itself usually contains many 'views' or 'presentations'; think of 3D scans of a mask, photographic pictures, X-ray data, textual descriptions . . . Many of these views themselves consist of several data objects: frontal pictures of Tutankhamen's mask, dating from 1799 up to 2007. Of course, we don't know in advance all possible future applications, but our data model should at least support the following applications:

1. *Show/navigate the data*
   We want to be able to present and access easily all parts of the data, from the most general overview to the smallest detail. In this example it means accessing data from airplane flyovers of the terrain up to the tiniest inscription, from old black and white pictures to X-ray data. And be only a 'click away' to navigate from e.g. this inscription to comments from scholars on it and translations in modern English.
2. *Present the data in an exhibition*
   We must be able to distill self-contained subsets of this information for expositions, textbooks, museums etc. For example, distill sufficient material for an on-line exhibition of pictures from the excavation of grave KV21 in 1817.
3. *Preserve the data*
   We want to maintain the information for at least 100 years. This not only means preserving the data in the literal sense of the word, but also keep it accessible under the fast changing technological conditions, and update it with any new information that might come available.

Apart from the previous example with large data collections, there exist also many small applications in the cultural heritage domain which require only a limited usage of a confined data set. Due to this diversification, it's not possible to define one single infrastructure technology, appropriate to support all cultural heritage applications. Therefore, we propose a general framework as an architecture, which can be tailored to any real application.

The requirements, as summarised in the next section, define the boundary conditions for a cultural heritage architecture. Section 3 explains why we prefer the information processing view over the pipeline view as the basis for a common infrastructure. Section 4 shows how data objects can be used to group all related information in an efficient way. These data objects are stored in an information management system, as described in section 5. The paper concludes by illustrating the concepts in two implementation examples.

2. Requirements

The important considerations of the research on stakeholder needs for the definition of a common infrastructure are:

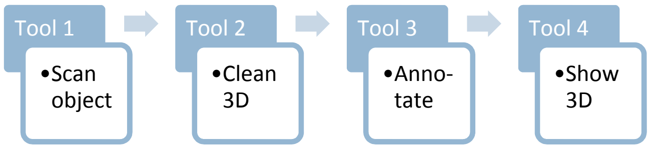
- need for a framework to guarantee interoperability among existing systems and facilitate data reuse;
- need for sustainability in regard to the accelerating progress of technology;
- need to maintain international standards;
- need to encourage an open source approach;
- attention to safeguarding scientific accuracy during data collection and processing by making a clear distinction between fact and hypothesis.

Good interpretation of cultural heritage also means that a wide range of sources (oral, written, research, traditions . . .) is used to create the interpretation and presentation, as the result of a multidisciplinary study. This not only means that the framework is conceptually valid for all cultural heritage domains but also that it allows exploiting all available information sources. In technical terms, this means that its data structures allow information from multiple domains. Today, databases in cultural heritage are usually tailored to a single discipline and are not suited to contain multiple domain information, created by a multitude of researchers, physically present at different locations.

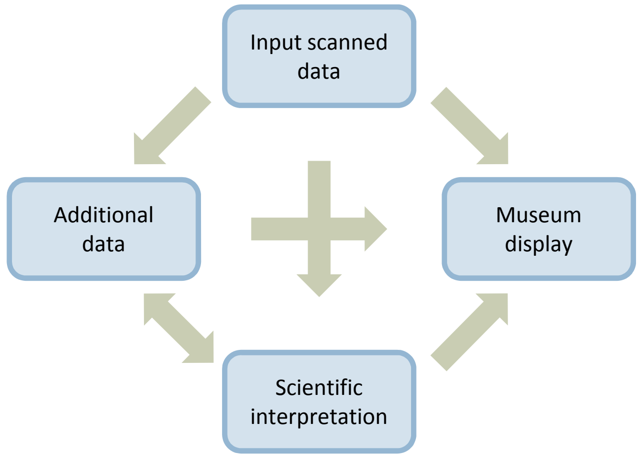
To ensure proper involvement of all parties concerned, interpretation technology needs to be anchored within the stakeholder communities, and not be the monopoly of high tech companies. In technical terms, we need an affordable and easy content management system as the kernel of interpretation systems, which must be designed to be adaptable and open.

From the technical point of view, two additional requirements are added. The framework must be as implementation technology independent as possible, since we want to prevent it from becoming outdated too quickly. This also implies that the framework cannot be restricted to the currently available tools and applications only. It should be able to embrace new emerging technologies.

To guarantee sustainability, the framework should be based on international standards as much as possible. The open source approach is also encouraged, but not enforced. Users should be able to use commercial tools and businesses should be able to develop based on a common infrastructure.



(a) Simple pipeline processing



(b) More complex application

Figure 1: Examples of processing in applications.

3. The concepts behind a common infrastructure

An application in general consists of a number of tools working on data. Each tool performs a well-defined subtask of the application. In practice the tools often form a sequence of operations: a tool works on the data generated by a previous tool and delivers its data to the next tool, as shown in figure 1a. This is commonly referred to as a "pipeline". On the other hand complex applications, such as the one in figure 1b can't be described as a pipeline: it consists of several sub-applications which combine and iterate upon exchanged data. This data can even be generated or used at different times and places.

3.1. Classification of tools and applications

There are several orthogonal ways to describe cultural heritage applications. A first dimension consists of the seven functional technology areas as described in the application for the Epoch network. These areas correspond to the numbered boxes in figure 2. This classifies cultural heritage activities from a computer science point of view.

A second dimension describes the chain of practice in archaeology and museums. This cultural heritage pipeline corresponds to the unnumbered boxes in figure 2. This figure also visualises which technology area impinges upon which part of the archaeological or cultural heritage process, but without the feedback loops and iterations. It is important to note that due to the loops and iterations the process is not a real pipeline where information enters on one side and propagates through transformations towards the end of the pipeline.

A third dimension enumerates the different tools for the different application domains. The domains correspond to (subsets of) the different stakeholder communities as described by the Stakeholder Needs team (Heritage Policy, Local Authorities, Cultural Heritage Sites, Museums and other Cultural Heritage Organisations, Associated Communities, Tourism, Education, and Technology). Tools may be specific for a domain or may be shared. For instance, avatars can be used in many applications but Harris matrix tools are very specific to excavations.

This classification is a good way to categorise tools, but it is less useful for defining a common infrastructure for cultural heritage applications. Such an application can be represented by a cloud of points, which cover only a part of this classification space. Even for applications with similar goals, their point clouds can be very dissimilar. On the one hand, this classification space is too extensive to derive a common infrastructure from it. On the other hand, each dimension in itself is not enough to describe all applications so it cannot be the sole basis of a common infrastructure. Therefore, we need another approach to tackle the problem.

3.2. The information processing view

Another way to look at the cultural heritage processing is to follow the information through the application. A first conclusion is that very few applications cover all aspects of data processing in a domain. Most applications deal with only one aspect or at most a few. A consequence of this fact is that applications are not self-contained: they use data from or produce data for other applications, so we need intermediate storage. This is certainly needed if loops and iterations are present, not inside but between applications. Furthermore, data is often re-used by different applications. A final conclusion is that many applications may take their inputs from multiple other applications.

For this kind of processing, the information flow can be depicted as given in figure 3. All the data is stored in the information store. This data can not only be used to generate output, but it can also be modified or synthesised into new data. Having a central point of information exchange has the advantage that it is easier to decouple

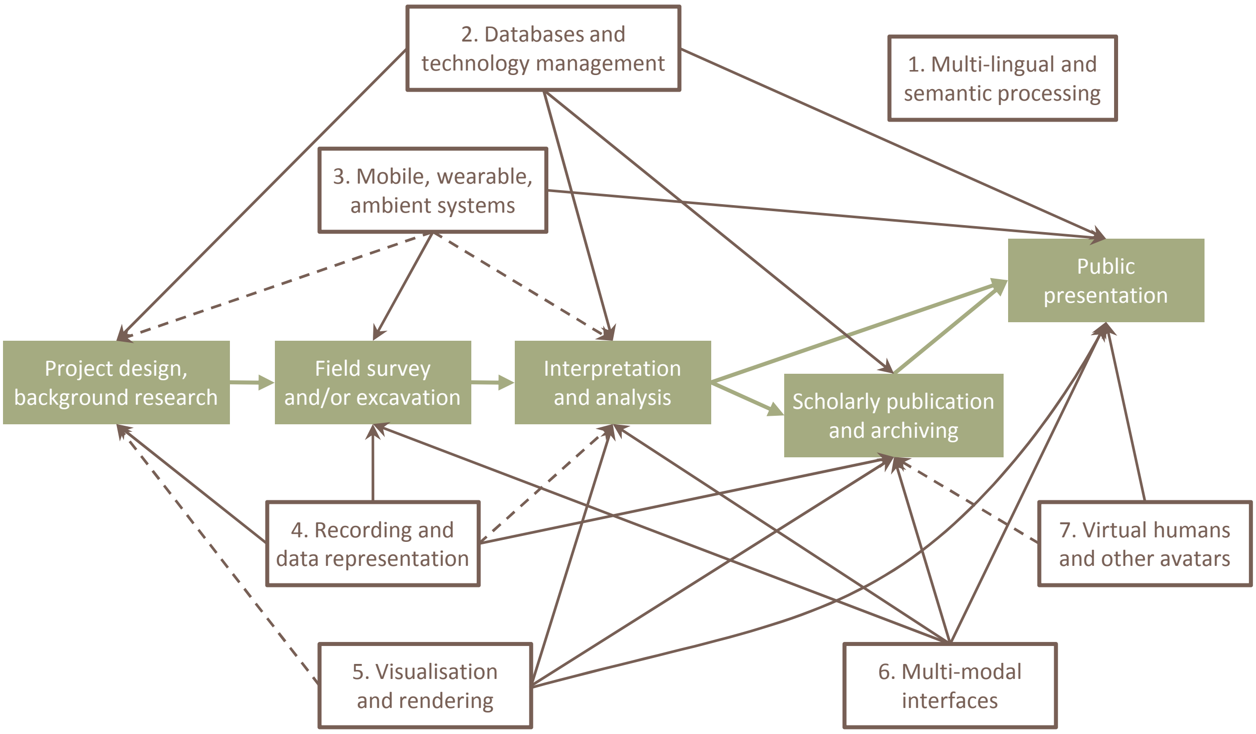


Figure 2: Relation between the typical cultural heritage pipeline (the filled boxes) and the functional areas (the numbered boxes) as proposed by Nick Ryan.

the generation, manipulation, and final use of information when needed, as requested by the stakeholders.

The central element is the information store. In theory, it stores all cultural heritage data, which was ever produced or which will ever be used by cultural heritage applications. In order to avoid mixing up measured data and interpretation, this store should hold the raw data (the original "mother data" as well as the transformed data) with the accompanying metadata, which describes all extra information and transformations applied to it.

Data is input in the data store in two ways. First of all, data producers generate data for it. Examples of such producers are tools to measure/scan objects and field data, or tools to add comments. Another way to generate data is to modify or synthesise existing data. The data synthesis implies that multiple sources are used to generate new data. Examples of these transformations are generation of virtual objects and worlds, annotations, and the interpretation of data. Finally, the data from the information store can be output to users

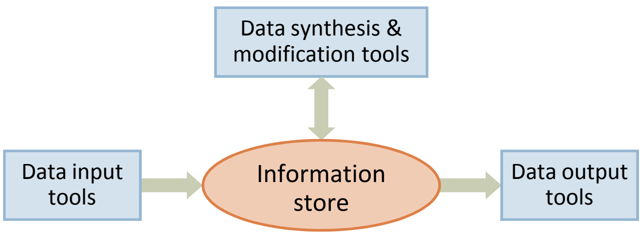


Figure 3: Information processing view of cultural heritage applications.

by data consumers. Usually they also use multiple inputs. Examples of data output are not only all kinds of display, with or without user interaction, but also the outcome of queries for information from the data store.

Of course, a real application cannot hold all cultural heritage data and all tools for working on it. In practice we only need a subset of the information store as well as of the tools, just enough to get the job done. If needed for other applications, we can merge it with other information stores. This is not a problem as long as all information stores are based on the same foundation and data synchronising tools are available.

3.3. Concepts of an architecture

As illustrated in the information processing view of the process (figure 3), there are three parts in solving this problem:

- First of all you need an appropriate set of tools to operate on the cultural heritage data. These tools produce data, ranging from text input over sound to video and 3D scanning; they consume data, e.g. for display, or they modify existing data, by enhancing and/or synthesising the data. Since many links exist between the different data items, we may need a way to keep all related items and their link information together as one data object. Section 4 proposes a way to achieve this.
- The way to exchange the data between the tools is through the information store. This information store does not necessarily correspond to a physical storage such as a database, but it may. For instance, when data is processed within an application in a single pipeline, the intermediate data can be kept in memory. But when different applications are using and exchanging information,

a more formally defined storage implementation may be required. Section 5 proposes an implementation of such a cultural heritage information store based on existing content management systems.
- Finally, exchange formats (shown by the arrows in figure 3) must be defined to guarantee a proper exchange of information, raw data as well as metadata. For long time preservation additional requirements should be met. Epoch is not proposing to define new or adapted standards. In our opinion, enough standards for our goals are available. Often there are even too many to choose from, so we rather propose guidelines to select an appropriate standard. Section 6 contains some agreed upon guidelines for selecting standards for multimedia and 3D.

4. Cultural heritage data

Cultural heritage data usually consists of different types of raw data (text as well as multimedia) and relationships between them. To keep all this related information together, we group it into a data item, which we'll call a "Cultural Heritage Data Object" (CHDO) from now on. An example of such a CHDO is a 3D representation of an object. The CHDO stores not only the 3D model itself, but also the way to obtain it (original scan data, information about processing steps and their parameters). Furthermore other annotations (e.g. in CIDOC-CRM format), other metadata (e.g. intellectual property rights) and links to other objects and descriptions can be provided.

Of course, a CHDO can not hold all the cultural heritage data of one application. For instance, all information on an archaeological site, ranging from 3D models to annotations and para-data, is too extensive to store in one object. For practical reasons CHDOs must have a finer granularity. As long as proper linking tools are provided, this is not a real limitation.

4.1. The diversity of cultural heritage data

In the different cultural heritage applications one finds quite a diversity of data characteristics. This is illustrated by the following three aspects: classes of typical usage, types of data, and the required accuracy. The CHDO should be able to accommodate all these different aspects.

Typical usage. The table below summarises the different requirements for the three different classes of usage: preservation, exchange, and presentation of the data.

	Preservation	Exchange	Presentation
Sustainability	important	less relevant	irrelevant
Standards used	few non-proprietary	few	application driven
Extend standard	avoid	allow	allow
Compactness	desirable	desirable	irrelevant
Fast running	irrelevant	less relevant	important
Fast development	irrelevant	less relevant	important
Wide use	irrelevant	desirable	desirable

An important aspect of cultural heritage data is its sustainability, but this is only important for preservation data, i.e., data which should be preserved for at least some decades. The life span of exchange data is much smaller and its format is more determined by the properties of the tools which are exchanging data. On the other hand, the life span of display or presentation data solely depends on the presentation application. Its storage format must be optimally tuned to the (speed) requirements of the application. Typical presentation standards, such as SMIL, are not very suited for preservation purposes.

The preservation format is usually considered to be the master format. This means that converters must be available to transform the preservation format into the exchange and the presentation format. Therefore, reusing the preservation format as exchange and/or presentation format reduces the conversion effort. Furthermore, one has to be aware of the fact that every conversion may lose information and/or accuracy. This is one of the reasons to avoid in-place processing of the preserved data.

Data types. Cultural heritage data usually has a heterogeneous nature. The tools work on different types of recorded data (images, 3D models, GIS information, text notes ...) as well as on different types of semantic data (meta-data and para-data, such as provenance, interpretation and other knowledge). For a peer review of the interpretation it is necessary to store all the information on the data processing steps, including programs and scripts used together with their parameters. Additionally, presentation applications often require user interaction, so they need to store animation and event processing.

A very important aspect of all these cultural heritage data is their relations, including the hierarchical aspects. Therefore the linking facilities must be rich enough to support all kinds of links, including links to a part of an object (e.g., the head of a 3D model of a statue).

Required accuracy. From the accuracy point of view, two different classes of data can be distinguished:

- Recorded data is recorded or generated data, which might be used later for analysis and interpretation. This data must be preserved with a maximal accuracy, i.e., the accuracy of the recording equipment.
- Illustrations (e.g. textures, movies, and sound) are only meant to look or feel or sound good. The accuracy is less important here as long as the perceptual quality is appropriate. This data can be preserved with a minimal accuracy corresponding to the required perceptual quality.

These classes have quite different requirements on the formats, e.g., on the allowed compression. Since recorded data must be kept at maximum resolution, compression of recorded data must be lossless. On the other hand, only the perceptual property of illustrations must be maintained so they can undergo a lossy compression.

4.2. Choosing a CHDO format

The CHDO is the basic unit for storing cultural heritage data. It should at least be defined for preservation purposes, although it can be useful for exchange of data also. It contains multiple types of data, ranging from recorded data to semantic information. It can be linked to other information (including other CHDOs and parts thereof) to form another CHDO at a higher hierarchical level.

If it is used for data preservation, the CHDO should be write-once. This means one can add information and links to it, but never change or remove information. Modifying a preservation CHDO should not be done in-place; it must result in a new CHDO.

If it is used for data exchange, the CHDO must accommodate multiple standards or exchange formats. If tools support a common exchange format, it is very inefficient to convert it to another intermediate format just for the sake of having only one exchange standard. Of course, it's recommended to have guidelines in order to keep the number of exchange formats as low as possible.

There are three ways to implement a CHDO:

- One can select one data type as the primary type, select a standard for it and extend the standard to accommodate other types of information. Let's call this the "*primary type implementation*". An example would be to choose 3D as the main data type, select X3D as the corresponding standard and extend X3D with additional fields to store other data, such as CIDOC-CRM data.
- One can choose an appropriate standard for each type of data and provide linking information between the elements. A CHDO element is a part of the CHDO data, which has a certain data type. The CHDO becomes a collection of loosely coupled information and the CHDO coherence is solely defined by the linking information, distributed over the elements. Let's call this the "*juxtaposition implementation*".
- One can choose an appropriate standard for each type of data and provide a tight coupling between the elements by using a container format. The CHDO is the container together with its elements. The container format contains all top level CHDO information as well as intra-CHDO linking information. Let's call this the "*container implementation*". Figure 4 illustrates the concept: each subtype has its own format and elements use the standard designed for them; the container holds the relations and all data not specific to one element. Note that this is only an illustration of the concept: not all types are present and elements may be used more than once (e.g., more than one 3D model).

The main disadvantage of the primary type implementation is that it abuses the primary type standard (X3D in the example above). It relies on the fact that this standard allows standard extensions (which is the case for X3D). The problem with extensions is that they are not part of the standard, which makes them unusable for preserving information. One could get round this problem by providing all

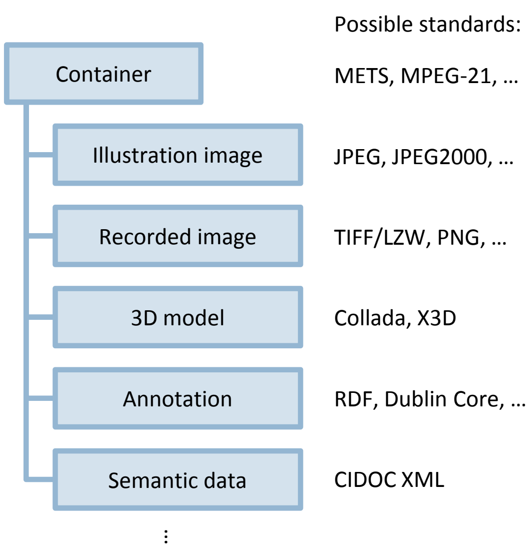


Figure 4: Using the container implementation to keep related data together.

the necessary tools to handle the extensions (e.g. an Epoch specific viewer) oneself, but this is very costly and very hard to sustain in the long run. A better solution is to have your extensions incorporated into the official standards by having them accepted by the international standard committees. Not only is this a very time consuming effort but it will only work if cultural heritage is a substantial part of the target group of the standard, which is not the case for X3D. So the primary type implementation can not be used as a preservation format.

The juxtaposition implementation solves this problem by using a specific standard for each type of data. However it still provides no solution for the relations between CHDO elements. Since these links are completely external to the CHDO, we not only need a fixed standardised linking procedure but each CHDO element must also have a registered identifier in order to guarantee sustainable links. Therefore the juxtaposition implementation is less suited as a preservation format.

Almost all of these problems are solved by using a container implementation: only extra-CHDO links have to be taken care of. But the container implementation has an important additional advantage: it decouples the standardisation aspects of all the different data types. With a container implementation we are not looking for a big global common infrastructure, but we divide it into elements, each with their own optimised implementation, and we look for a way to combine them into one CHDO ("divide et impera" as the Romans told us). This not only makes the problem manageable but it allows us to switch a standard implementation for one type without impact on the rest of the CHDO. This flexibility is certainly an advantage for exchange formats: we can replace JPEG by JPEG2000 without redefining the CHDO if all tools support it.

4.3. Requirements for a container class

Obviously, the formal design and choice of a container format is something that needs planning and studying. The general needs are:

- An existing standard should be preferred whenever possible. If that is not possible, extensions to existing standards or profiling the standard should be preferred over inventing something completely new.
- Care should be taken to choose the simplest solution that serves the requirements. Container classes can be very involved.
- Flexibility for accommodating future changes, because of unforeseen future requirements, should be possible.
- The standard should be believed to be stable and in use for a sufficiently long time.
- Intellectual property rights (IPR) issues should be incorporated. Not only for the whole data object, but also for each of its data items separately.

The technical needs are:

- Preferable it is based on XML. As it is W3C recommended and in general use, XML looks like the preferred way.
- Registration of unique identifiers to items or links should be possible. Think of e.g. DOI [DOI]. There are alternatives though.
- Immutable objects: data objects should probably be write-once. As they might be referenced from external sources, any change in them might invalidate these links. The only way to prevent this (apart from unpractical back-links), is not to allow changes to objects at all. Modifications to an object will create a new (version of the) object.

4.4. Existing standards for container classes

Quite some container class standards exist, but most of them are very much tuned to specific applications. Here we need containers suitable for "all multimedia" in the wide sense of the word. That restricts our choice to two standards specifically developed for this purpose, and one very general solution.

- *DIDL from MPEG-21* [DIDL]
  MPEG-21 standardises a multimedia framework. Of interest to the cultural heritage community, as far as is clear for the moment, is mainly Part 2, DIDL. As is obvious, the terminology is tuned for multimedia objects. Nevertheless, the possibilities of the framework are much wider.
- *XPackage and RDF* [RDF, XPKG]
  They can describe various resources and their associations. The major difference with DIDL is that RDF is less specific, less tuned to a particular application, but therefore also more flexible for non-standard needs.
- *METS* [METS]
  METS is a framework to structure relevant metadata (descriptive, administrative and structural) in the digital library world. It is increasingly used in digital libraries and museums.

At least each of MPEG-21/DIDL, RDF and METS can do the job. However, there is more active development in METS and more interaction between METS and other important standards in this field. So at this moment, we prefer METS as the import/export container format.

5. Cultural heritage information management

The previous section defined the CHDO to keep all related information together. However in practical implementations, it's impossible to store all data inside one CHDO. And even if it is possible, it may not be desirable for most applications, which access only small portions of the entire data set. So the actual granularity of the CHDOs in an application depends on the data usage as well as on implementation restrictions. This means that in most implementations we end up with a collection of CHDOs and relations between them. This is illustrated in figure 5.

For small applications one should be able to dispose of all overhead including the information management system. Applications which only manipulate semantic information will be implemented using dedicated software and database. It's clearly overkill to transform the semantic network into a network of CHDOs. This implies that the choice of a information management system should not dictate the formats and standards used to store the raw data, nor should it put restrictions on the tools needed to search or visualise the data.
But if one at a later stage wants to connect the semantic information to other information, it should be possible through an information management system. As illustrated in figure 5, this can be done by considering the semantic network as one data object or by splitting up the network in relevant parts, which can become part of other data objects.

In fact, all this is nothing else than the information management systems in the Digital Library approach. However, there are some specificities compared to a classical digital library system; rules on data items have to be much less rigid: new data can be added to old items, new types of data will appear, classification systems and interlinking can change during the lifetime of the data set and novel, unforeseen use of the data set cannot be excluded. Nevertheless, it would be nice if this approach could leverage on what has been developed in the digital library community before. Moreover, there have been ample discussions on long term aspects of storage, exchange and accessibility, which do need to be reassessed in the specific case of cultural heritage, but are largely identical. On top of that, there is existing software and freeware available, so implementation and maintenance costs can be minimised.

An extensive search for software that would implement a suitable modular content management system was done, and returned a set of tools that were usable and suitable for a test scenario, but none that was exactly tailored for Epoch-like environments. DSpace [DSP] is frequently used by libraries and archives and fully supports METS and MPEG-21/DIDL. Fedora [FED], not to be mixed up with the Fedora Linux distribution from RedHat, seems a bit more suitable for our purposes because of its 'dissemination objects' (kind of filter plug-ins in the server) and fine grained access control, but might have a smaller community than DSpace. Internally, recent versions of Fedora use FOXML (Fedora Object XML) for performance reasons. FOXML objects can be exchanged with other Fedora repositories, but import and export in METS is supported; METS1.4 and MPEG-21/DIDL plug-ins are under development.

6. Suggested standards and formats

Since Epoch is not a standardisation body, we provide only guidelines for selecting standards ("best practices") for intra-Epoch exchange (limited number of formats) and preservation in domains with Epoch interest. In our context, standards are about storing/exchanging information, not about visualisation.

Standards for audio, vector graphics, 2D images, and video were proposed, both lossless and lossy. More details can be found on the Epoch web site [MMS]. For 3D data the use of Collada is proposed. In Epoch a "Collada light" is used plus binary object representations. U3D seems very well suited for storing compressed 3D data.

Standards are frozen, so you have to switch to a new standard if

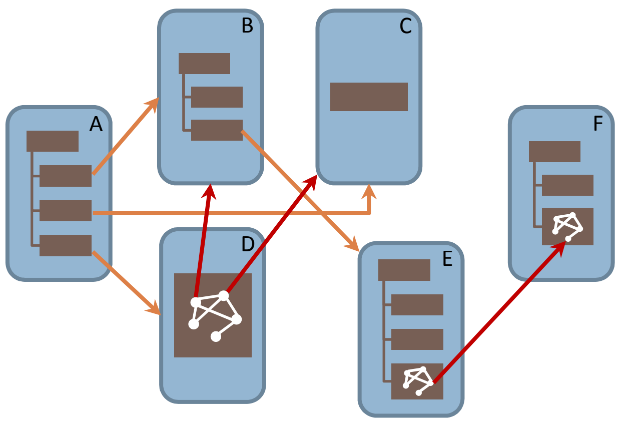


Figure 5: Linking data objects. The objects can be simple objects (C or the semantic networks D), or container objects. Links can represent an hierarchy (as from A to B, C and D) or direct relationships (between B and E). From inside an element links are not only possible to other objects (e.g., from D to B) but also to elements inside objects (e.g., from E to F).

you need new technology. Therefore, guidelines will evolve, certainly for exchange purposes.

7. Implementation examples

The infrastructure must be validated against real cultural heritage processes. This section describes two chains of tools which are not only used to demonstrate how to use the tools in an application, but also to test the interaction of the tools and the exchange of data between them. These tool chains were demonstrated at the exhibition part of Epoch's final 'Rome Event'.

A virtual visitor centre. This tool chain allows people to virtually visit sites which are normally closed to the public or which no longer exist. (The demo uses Pompeii as an example.) The visitor is guided through the virtual site by an avatar. At predefined places the visitor can examine objects (e.g. buildings or excavated objects) using a separate 3D-viewer. The visitor can also ask questions about the objects, preferably in his natural language.

Fedora is used as the content management system of this application. Different types of data are available about the 'House of Pansa', such as descriptions, multimedia data and metadata. This information is available on-line to the user (on a web site or on a mobile device), but a poster for the museum is also present (cf. figure 6). All presentations come from the same data source, but are tuned to the specificity of the users, their needs as well as their equipment hardware, with possibly minimal manual interventions needed from the person editing and composing the presentations.

Archaeological field survey. This tool chain addresses both ends of the cultural heritage chain of practice, from data collection to public presentation. During field walking, an archaeologist finds a small and interesting object. Information from MobiComp [MOBI] and 3D-scanning is combined with hand-written notes and stored in Fedora. Actors carrying a multimedia guide access the model when appropriate, according to their profile and preferences.

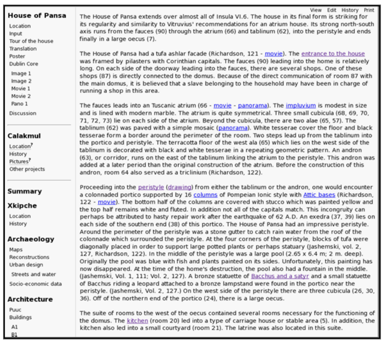
The infrastructure is based on Fedora and MobiComp. The models are stored in Fedora, while context elements are stored and handled by MobiComp. Contextual data in MobiComp is used to generate metadata for object models stored in Fedora. The models and metadata stored in Fedora are used by tools to generate context-dependent multimedia guides.

8. Conclusion

This paper shows that in spite of the diversity of cultural heritage applications some general concepts and guidelines can be established to provide interoperability of tools and sustainability of data. These concepts are not bound to the current state of available implementations, so they anticipate new technologies and solutions.

Acknowledgements

The research described in this paper were funded by the Epoch project and the Interdisciplinary Institute for Broadband Technology (IBBT). The concepts concepts described are the result of numerous discussions, by email as well as in meetings, between the partners of activity 3 of the Epoch project.



(a) Web format

(b) Poster format

Figure 6: Different presentation results starting from the same data of the House of Pansa in the Fedora repository.

References

[DIDL] J. Bekaert, P. Hochstenbach, H. Van de Sompel: Using MPEG-21 DIDL to Represent Complex Digital Objects in the Los Alamos National Laboratory Digital Library. *D-Lib Magazine* 9,11 (November 2003). http://www.dlib.org/dlib/november03/bekaert/11bekaert.html

[DOI] The DOI System. http://www.doi.org/

[DSP] DSpace. http://www.dspace.org/

[FED] Fedora Commons. http://www.fedora-commons.org/

[METS] Metadata Encoding and Transmission Standard. http://www.loc.gov/standards/mets/

[MMS] Multimedia formats in Epoch. http://partners.epoch-net.org/common_infrastructure/wiki/index.php/Multimedia_Proposal

[MOBI] MobiComp. http://www.mobicomp.org/

[RDF] Resource Description Framework. http://www.w3c.org/RDF/

[XPKG] XPackage. http://www.xpackage.org/