Ontology-Driven Digital Preservation of Interactive Multimedia Performances

Alexander Mikroyannidis,[1] Bee Ong,[1] Kia Ng,[1] and David Giaretta[2]
[1]*ICSRiM - University of Leeds,*
School of Computing & School of Music, Leeds LS2 9JT, UK
[2]*STFC, Rutherford Appleton Laboratory, Oxfordshire OX11 0QX, UK*
caspar@icsrim.org.uk   www.icsrim.org.uk/caspar

Abstract

Interactive multimedia performances are rapidly gaining ground within performing arts communities nowadays, mainly due to breakthroughs in human-computer interaction technologies, such as human motion capture and analysis. This has brought forward the issue of digital preservation of these performances, so that they can be reconstructed in the future. This paper presents an ontology-driven approach to the digital preservation of interactive multimedia performances. An ontology model is proposed for describing the complex relationships amongst different components of a performance, as well as their temporal aspects and evolution over time.

1. Introduction

An Interactive Multimedia Performance (IMP) involves one or more performers who interact with a computer based multimedia system making use of multimedia contents that may be prepared as well as generated in real-time including music, manipulated sound, animation, video, graphics, etc.

An example of an IMP process is the one adopted in the MvM (Music via Motion) interactive performance system, which produces music by capturing user motions [1]. The system captures user motions using motion capture devices and stores them in a 3D format. These motions are then mapped into music by using a mapping strategy, with parameters provided through a GUI. The motion-music map is forwarded to the generation component which produces multimedia content.

IMP preservation is a challenging issue. In addition to the output multimedia contents, related digital contents such as mapping strategies, processing software and intermediate data created during the production process (e.g. data translated from "signals" captured) have to be preserved, together with all the configuration, setting of the software, changes (and time), etc. The most challenging problem is to preserve the knowledge about the logical and temporal relationships amongst individual components so that they can be properly assembled into a performance during the reconstruction process of the original IMP.

The preservation of IMPs produced by the MvM system comprises part of the Contemporary Arts testbed dealing with preservation of artistic contents, which is one of the three testbeds of the EU project CASPAR (www.casparpreserves.eu). The other two are Scientific and Cultural testbeds which are focused on very high volume and complex scientific data objects and virtual cultural digital objects respectively.

This paper introduces an ontology approach to describing IMPs for their preservation. A set of extensions for the CIDOC-CRM standard [2-4] are proposed, together with an ontology model for describing temporal facts. The remainder of this paper is organized as follows: Section 2 presents some metadata approaches to digital preservation. The applicability of CIDOC-CRM in digital preservation is discussed in section 3. Section 4 introduces the proposed extensions to CIDOC-CRM and section 5 presents the ontology model for the temporal enrichment of metadata. Finally, the paper is concluded and some plans for future work are provided.

2. Related Work

Metadata and ontologies have been proven an important factor in digital preservation. Metadata element sets designed specifically for preservation purposes include those developed by RLG Working

Group on Preservation Issues of Metadata (RLG) [5], CURL Exemplars in Digital Archives (CEDARS) project (www.leeds.ac.uk/cedars) [6], the metadata of the National Library of Australia (NLA) [7] and the Networked European Deposit Library (NEDLIB) [8]. A consensus effort was carried out by the OCLC/RLG Working Group on Preservation Metadata to develop a common metadata framework to support the preservation of digital objects, which was based largely on CEDARS, NEDLIB and NLA element sets [9]. The Preservation Metadata Implementation Strategies (PREMIS) Working Group later built on this framework a PREMIS data model and a data dictionary for preservation metadata [10].

The CIDOC Conceptual Reference Model (CRM) has been proposed as a standard ontology for enabling interoperability amongst digital libraries [2-4]. CIDOC-CRM defines a core set concepts for physical as well as temporal entities. This is very important for describing temporal dependencies amongst different objects in a preservation archive. A combination of core concepts defined in CIDOC-CRM and multimedia content specific concepts of MPEG-7 for describing multimedia objects in museums has also been introduced. A harmonisation effort has also been carried out to align the Functional Requirements for Bibliographic Records (FRBR) [11] to CIDOC-CRM for describing artistic contents. The result is an object oriented version of FRBR, named FRBRoo [12].

3. CIDOC-CRM for Digital Preservation

CIDOC-CRM was originally designed to describe cultural heritage collections in museum archives. The meta-schema of CIDOC-CRM is illustrated in Fig. 1. CIDOC-CRM's conceptualisation of the past is centred on Temporal Entities (e.g. events). People (Actors) and objects (Conceptual Objects and Physical Objects) involved, time (Time-Spans) and Places are documented via their relationships with the Temporal Entities.

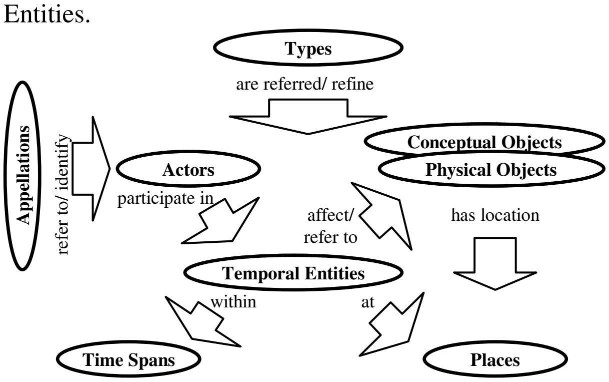

Figure 1. The meta-schema of CIDOC-CRM [4]

The CIDOC-CRM vocabulary can be used to describe a performance at a high level. However, more specialised vocabularies are necessary for the interactive performing art domain to precisely describe the relationship amongst the elements of a performance. For example, it is necessary to model how equipments are connected together in the performance. Some concepts representing digital objects have been very recently introduced in CIDOC-CRM for digital preservation purposes. Nevertheless, there is a need for documenting the relationships amongst software applications, data, and operating systems, as well as the operations performed on them. In addition, CIDOC-CRM is designed primarily for the documentation of what has happened, whereas in digital preservation, it is also required to document the reconstruction of a past event from preserved components.

4. CIDOC-CRM Extensions for Preservation of IMPs

In order to address the preservation of IMPs, we propose a set of extensions to CIDOC-CRM. These extensions have the following objectives:

- To provide a domain specific vocabulary for describing objects related with IMPs.
- To provide a vocabulary for describing the interrelationships between digital objects and the operations performed on them in the digital preservation context.

Fig. 2 shows the set of concepts that we have introduced in CIDOC-CRM to describe IMP objects. The extended concepts are prefixed by IMP and an identification number. The original CIDOC-CRM entity and property names are prefixed by E and P respectively.

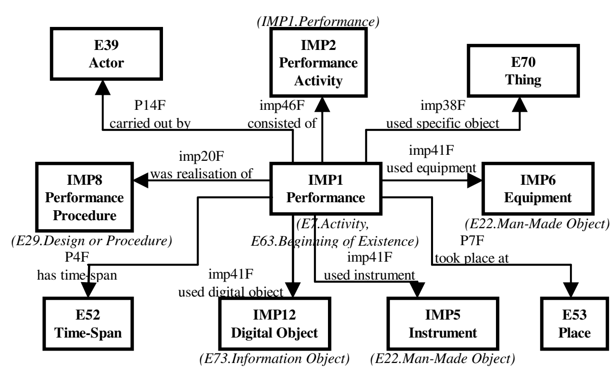

Figure 2. CIDOC-CRM extensions for describing IMP objects

More specifically, the following concepts have been introduced:

- *"IMP2.Performance Activity"*: for describing activities of a performance.
- *"IMP5.Instrument"*: a specialisation of CIDOC-CRM "E22.Man-Made Object" for modelling musical instruments (e.g. cellos, violins, drums, etc.) used in a performance.
- *"IMP6.Equipment"*: a specialisation of CIDOC-CRM "E22.Man-Made Object" for modelling equipment (e.g. a microphone, a sound mixer or a computer, etc.) used in a performance.
- *"IMP8.Performance Procedure"*: a specialisation of CIDOC-CRM "E29.Design or Procedure" for describing the procedure in which a performance should be carried out.
- *"IMP12.Digital Object"*: a specialisation of "E73.Information Object" for describing digital objects.

As shown in Fig. 3, "IMP12.Digital Object" has two subclasses: "IMP17.Digital Data Container" and "IMP18.Digital Data Object". A digital data container ("IMP17.Digital Data Container") is a container of one or more digital data objects ("IMP18.Digital Data Object"). An example of digital data container is a file. The bit stream contained within the file is considered as a digital data object. This separation is necessary to model a bit stream in memory or in cases where multiple bit streams carrying different information carried by a single digital data container. A special type of digital data object is a computer program (IMP13.Computer Program). In this case, the bit stream is a set of instructions to be executed by a computer. There are two specialisations of computer programs: "IMP14.Operating System" and "IMP15.Software Application".

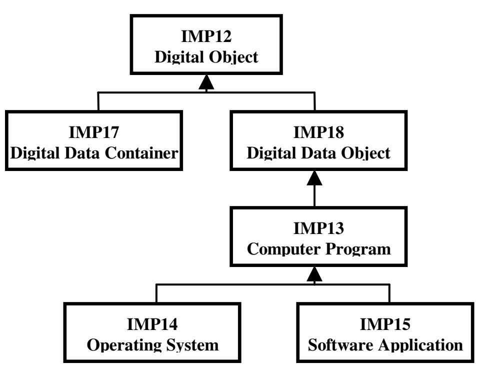

Figure 3. Classification of digital objects

Operations on digital objects can be described using "IMP26.Digital Object Operation", which is a specialisation of CIDOC-CRM "E5.Event". A number of subclasses of "IMP26.Digital Object Operation" have also been defined to deal with common operations such as creation, duplication, transformation, modification, access and deletion. This is necessary in the preservation context, where the history of a digital object needs to be documented.

5. Bitemporal Ontology Modelling

In order to capture temporal facts in our preservation ontologies, we propose the use of the bitemporal ontology model of Heraclitus II [13]. The Heraclitus II framework considers ontologies as a semantically rich knowledge base for information management and proposes ways for the management and evolution of this knowledge base. Heraclitus II uses an ontology model that is based on the object model defined by the Object Data Management Group (ODMG) [14] and more specifically on TAU [15]. The TAU model is an extended version of ODMG that supports modelling and reasoning about time and evolution.

Ontology modelling in Heraclitus II is bitemporal, allowing for ontology representation over two dimensions of time: *valid* and *transaction* time. The valid time of a fact is defined as the time when that fact is true in the modelled reality. The transaction time of a fact is defined as the time when that fact is current in the knowledge base and may be retrieved. Valid times can belong in the past, present or future and are usually supplied by the ontology author. Transaction times are provided by the ontology management system, cannot change and are bounded between the knowledge base creation time and the current transaction time.

Ontology objects, namely concepts, relations and instances, can be associated with transaction time, valid time, both (bitemporal), or none (static). This modelling allows for *retro-active* as well as *pro-active* changes to be captured and represented on the knowledge base. A retro-active change occurs when a fact that is entered at a certain transaction time in the knowledge base, has been valid in the real world before this transaction time. On the other hand, when the valid time of a fact is greater than its transaction time, then a pro-active change is captured in the knowledge base.

In digital preservation, certain changes have to be monitored in order to keep the archived IMP up-to-date and be able to reconstruct it at any time. These changes mainly regard the hardware and software the IMP was produced with, as well as changes in the environment of the IMP, such as changes in the performers' behaviour or in the environment setting of the performance. Modelling these aspects in a bitemporal ontological form can help capture and document any occurring changes more systematically and efficiently.

For example, the behaviour of the performers is an essential factor for maintaining the authenticity of an IMP. It is very likely though, that changes will occur over time in the fashion a performer plays an instrument, thus compromising the authenticity of the reconstructed performance. One way to address this issue can be modelling the original behaviour of the performer within an ontology, with the use of bitemporal concepts and relations between them. We can then capture any changes in the performer's behaviour and represent them on the corresponding bitemporal ontology objects, thus keeping the evolution history of the performer's behaviour in our knowledge base. In this way, we should be able to re-interpret the performer's behaviour in order to adapt the rest of the IMP accordingly.

6. Conclusions and Future Work

The present paper explored the area of digital preservation of IMPs. An ontology-driven approach has been proposed, by extending the current concepts defined in the CIDOC-CRM standard, for preservation of IMPs. A number of concepts describing the performing art domain, as well as digital objects have been proposed. In addition, a bitemporal ontology model has been presented, addressing the temporal aspect of performances and their evolution over time.

As future work, the authors are planning to evaluate the proposed CIDOC-CRM extensions using MvM performance data. The proposed ontology will also be integrated with the architecture of CASPAR project for use by its software components.

7. Acknowledgements

Work partially supported by European Community under the Information Society Technologies (IST) programme of the 6th FP for RTD - project CASPAR. The authors are solely responsible for the content of this paper. It does not represent the opinion of the European Community, and the European Community is not responsible for any use that might be made of data appearing therein.

8. References

[1] K. C. Ng, "Music via Motion: Transdomain Mapping of Motion and Sound for Interactive Performances," *Proceedings of the IEEE*, vol. 92, 2004.

[2] M. Doerr, "Increasing the Power of Semantic Interoperability for the European Library," in *ERCIM News*, vol. 66, 2006.

[3] T. Gill, "Building semantic bridges between museums, libraries and archives: The CIDOC Conceptual Reference Model," *First Monday*, vol. 9, 2004.

[4] M. Doerr, "The CIDOC CRM - an Ontological Approach to Semantic Interoperability of Metadata," *AI Magazine*, vol. 24, 2003.

[5] B. Berger, J. Coleman, W. Cromwell-Kessler, R. Dale, B. DeCandido, C. Montori, and S. Ross, "RLG Working Group on Preservation Issues of Metadata - Final Report," RLG and Preservation 1998.

[6] M. Day, "Metadata for Preservation - CEDARS Project Document AIW01," UKOLN 1998.

[7] National Library of Australia, "Preservation Metadata for Digital Collections," 1999.

[8] C. Lupovici and J. Masanes, "Metadata for long term-preservation," NEDLIB 2000.

[9] OCLC/RLG Working Group on Preservation Metadata, "Preservation Metadata and the OAIS Information Model - A metadata framework to support the Preservation of digital objects," 2002.

[10] PREMIS, "Data Dictionary for Preservation Metadata - Final Report of the PREMIS Working Group," OCLC and RLG 2005.

[11] FRBR, "Functional Requirements for Bibliographic Records - Final Report," International Federation of Library Associations and Institutions (IFLA), Frankfurt am Main, Germany 1997.

[12] M. Doerr and P. LeBoeuf, "FRBRoo Introduction," 2006.

[13] A. Mikroyannidis and B. Theodoulidis, "Heraclitus II: A Framework for Ontology Management and Evolution," in *Proc. 2006 IEEE /WIC/ACM International Conference on Web Intelligence (WI 2006)*, Hong Kong, China, 2006, pp. 514-521.

[14] R. G. G. Cattell and T. Atwood, "The Object Database Standard: Odmg-93: Release 1.2." San Francisco: Morgan Kaufmann Pub, 1996.

[15] I. Kakoudakis and B. Theodoulidis, "TAU: Towards a Temporal Object-Oriented DBMS," in *Proc. 9th International Database Conference (IDC)*, Hong Kong, 1999.