The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST (2006)
M. Ioannides, D. Arnold, F. Niccolucci, K. Mania (Editors)

Documenting Events in Metadata

M. Doerr[1] and A.Kritsotaki[1]

[1]ICS-FORTH, Heraklion, Greece
{martin,,athinak}@ics.forth.gr

Abstract

In this paper we outline the importance of event-centric documentation for structuring cultural metadata and historical context. Historical analysis can be seen as an analysis of events involving participation of people and things, meeting each other and thus creating history. Event modeling is so abstract that it can be used to describe cultural items and documentations of scientific observations. This work aims to show how event modeling provides more accurate information about life histories, relates and aggregates relevant information, and so helps to a more effective search and retrieval than currently achieved with Dublin Core and VRA.

Categories and Subject Descriptors (according to ACM CCS): J.2 [Computer Applications]: Archaeology

1. Introduction

Event modeling is a major aspect for cultural – historical analysis because it is an essential part of the complex knowledge required for historical and cultural information; unfortunately, very few approaches focus on event documentation about cultural objects. They usually focus on detailed documentation about the objects and their particular features.

Documentation is an interpretation of cultural materials in relation to a historical context, which can be described in terms of events and processes. Historical context can be abstracted as things, people and ideas meeting in space-time [SAC*06]. For example, a type of an artifact, a style, results from a production event. Historical analysis can be seen as an analysis of events involving occurrences of agents/participants, presence of people and things (material or immaterial), meeting each other and thus creating history as a "network".

The abstraction of all the different kinds of events into simple meetings is a very powerful simplification for core documentation of cultural items and documentations of scientific observations.

Event-centric documentation provides a more accurate view of the past or current life history of a cultural object. Focus on factual information representation in contrast to categorization interprets more effectively history and especially, heterogeneous and complex information resources that are lost, not accurate or unrelated and need to be linked and interpreted in order to capture knowledge. It provides access to information about research and interpretation of the past, relates information sources and helps to a more effective search and retrieval.

Modeling of events can be used for the representation of metadata and content relationships as well, such as participation in an event, part-whole relation, reference information and classification [DIL*06] which are the most fundamental relationships that connect things, concepts, people, time and place. Modeling changes of state (based on criteria such as when, where, who etc.) provides more accurate information about life histories and also relating and aggregating relevant information and knowledge.

Even a description of cultural material is an observation which can be documented as an event. Events enable the construction of related information networks about history of things from the past.

2. Related Work

Event –modeling and documentation is not a common practice for the majority of the standards used in cultural documentation.
Only CIDOC CRM [Doe03] proposes a structure based on documentation of events and processes.

CIDOC CRM (ISO/FDIS 21127) is a standard for the semantic integration of cultural information. CIDOC CRM develops a general ontology [Gua98] about cultural documentation. It doesn't define terms (vocabularies) but relationships between entities. It is a model of 80 classes and 130 relationships, suitable to capture the underlying semantics and metadata of cultural documentation. It is based on the modeling of events and so it can be used both for the representation of metadata and complex content summarization as well. Its approach to event modeling is simple, generic and abstract in order to describe not only cultural materials but also scientific observations.
CRM uses four fundamental principles:

1. Participation in an event (e.g. creator, contributor, publisher, birth date, birth place, creation date, place of find, designer, project leader etc.
2. Part-whole relation.
3. Reference (e.g. subject, “aboutness”, representation)
4. Classification.

The basic idea is that historical context can be represented by things, people and ideas meeting in space-time. CRM proposes a simple schema for summarization of historical facts. The past is formulated as events involving “persistent items”, presence of things creating, in this way, a history of lifelines of things (meeting in discrete events). This general principle based on events definition can be used to model a variety of relationships.

It is a model which emphasizes on relationships rather than individual concepts or vocabularies; CIDOC CRM is an ontology, which allows for creating global networks of related knowledge.

CRM Core, on the other hand, is a metadata schema. It differs from CIDOC CRM in the following respect: CRM Core is a unit of documentation dedicated to a description of a specific item and not a semantic network of correlated knowledge; it is not ontology. It is made so that information from multiple instances of CRM Core about diverse items can be merged univocally into a knowledge network which instantiates CIDOC CRM. In other words, it is a means to manage the knowledge in the units in which it is produced by the experts.

The VRA Core [VRA02] standard provides a set of elements to describe works (inter alia, objects of material culture) of visual culture (and images that document them). It also defines vocabularies used for annotation. However, it fails to capture complex contexts of creation, use and generally, events and relationships (resulting from events), because information related to event context such as date, place and participants are disassociated.

The same practice is used in CCO - Cataloguing Cultural Objects [CCO05]. CCO is a guide used to describe cultural works and their images. It focuses on data content standards with emphasis on descriptive metadata. It relates, in a selective way, elements sets from VRA Core and Categories for the Description of Works of Art (CDWA).

Dublin Core [WKL*98] on the other hand, is a metadata standard, which defines a limited set of elements to describe general resources. It fails to capture complex historical material and context. It can not describe relationships, processes or phases, such as observations or research activities that can be related to a cultural object.

All the above indicated are schemata and not ontologies.

It appears that most standards focus on modeling categorical data in order to describe individual concepts rather than relationships. However, this practice cannot integrate or connect rich historical information.

3. Events as meetings

Ontologies describe possible *state of affairs*, a specific distribution of *potentially* observable *items*, i.e. material items, conceptual items and events, as well as their associated *relations* and *qualities*, over space and time [DPK*04]. Events can be seen as *particular states of affairs*, in which historical and archaeological phenomena are connected as a network of *persistent items* that meet in space and time.

Events consist of interactions [JLT*05] of participants, consist of “meetings”. *Meetings* are interactions of living or dead items that bring about changes of state.

History is a sequence of meetings. An event may cause or be caused by another event. Events order provides relative chronology by a relative order of creation and destruction events of participants (such as strata, finds, buildings etc.). These entities were present (“participated in”) at those events (deposition events, historical/ archaeological/architectural events - events of use and production, events- processes of information exchange). Primary evidence for the existence of past events are either their products, permanent traces, placement of objects or reports in written or oral historical records (information). Even immaterial items are regarded to participate in event via their carrier that necessarily reside on, such as human mind, paper, rock carving etc. - see fig.1 :transfer of information [DPK*04].

The action of observing/describing an event is part of the event (a meeting). Events are processes relevant to each other; specifically, are *non-instantaneous, finite* processes of a potentially *complex* nature.

Events cover the reality of archaeological evidence appropriately from the ontological, epistemological and mathematical point of view.

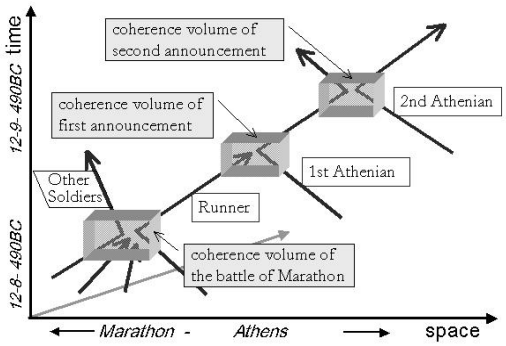

Figure 1: *Information exchange: Marathon runner “carries” a message*

4. Example

Examples implementing DC, VRA and CRM Core schema show different approaches in representing the required, relevant historical information. Our aim is to prove that examples that are not based on event documentation may yield wrong or insufficient conclusions during information search and retrieval.

“Monument to Balzac” is a characteristic example [Tan76]. It was commissioned to Rodin in order to honor

one of France's greatest novelists. Rodin spent seven years preparing for "Monument to Balzac" on several preparatory studies (showing different versions of Balzac). The final version (in plaster) was exhibited in Paris in 1898 (and it was then rejected by the conservative critics as an unfinished sketch). Only years after Rodin's death, his "Balzac" was cast in bronze (this is not an unusual practice; some works were even casts of early works that the artist never executed in bronze).

So, here we have a time-series (fig.3) presenting a construction of a work of art, "Monument to Balzac". It is a production event, a "meeting" based on our knowledge of a monument that was created. The "meeting" of the producer "Rodin" and his work "Balzac" happened in 1898 in France. Participation and presence is represented by the superproperty "P12 occurred in the presence of" which summarizes the roles of the participation of the actor and a thing, such as the role of a "producer" in case of Rodin and his product, "Monument to Balzac".

Since, biography (artist's dates) and sometimes locus of activity is useful information required to the art-historians, we also keep details about Rodin's life, such as when he was born, when he died, etc. This information can be represented in details by a birth and a death event.

Although, this biographical information about the artist seems to be unlinked to the production event, in fact it is related to the work and the date of the creation (independent descriptions can be part of the same event or linked through event description).

Information becomes more complex when it is required to represent our knowledge of the post humus bronze casting of the "Monument". This can be modeled as another production event, which continued the original production event of the work (a time-line for a production process) and occurred after Rodin's death (event). If we do not model this link/network of events and we attempt to search information about a post humus Rodin's work, we will probably find wrong information.

The same example implementing DC, VRA and CCO (fig.4,2), fails to show all the required related information because date, place and participants are described separately and are not related through their participation in discrete events. They can not show the relation between creator, date of creation, place and the object (which was created on a specific date and place, by a specific actor having a specific role and using a particular material). This approach fails to describe a history of processes/activities related to the cultural item.

Even structural and name changes, such as those of "Creator" in VRA Core 3.0 into the more generic term "Agent" in VRA Core 4.0 Version, can not solve the problem; (still, there is no connection to an event description). Moreover, they are characterized by inconsistency in proposing categories: for example, VRA Core 3.0 includes "Location. Discovery Site" and doesn't correspondingly include a "Date.Discovered".

Example (CCO)

Record type = item **Class**= sculpture
Work type = statue
Title = Monument to Balzac
Material and Techniques= bronze, plaster
Creator Display= Auguste Rodin (French, 1840-1917); Rudier (Vve Alexis) et Fils.
Role [link]=sculptor [link]: Rodin, Auguste
Role [link]=casters [link]: Rudier (Vve Alexis) et Fils
Creation Date = designed and produced in 1898, cast in 1925
Qualifier: design **Start**: 1898 **End**: 1898
Qualifier: casting **Start:** 1925 **End**: 1925
Subject = Balzac
Culture= French
Description = Commissioned to honor one of France's greatest novelists, Rodin spent seven years preparing for *Monument to Balzac.* When the plaster original was exhibited in Paris in 1898, it was widely attacked. Rodin retired the plaster model to his home in the Paris suburbs. It was not cast in bronze until years after his death.

Figure 2: *"Monument to Balzac" implementing CCO*

5. Conclusions

In this work we emphasize the importance of event modeling for historical analysis. Structures that are not based on event documentation fail to support meaningful information integration.

So, we propose a new metadata schema (CRM Core), which has comparable complexity with DC, VRA and higher generality; however it is capable to capture knowledge networks. It has the power to provide more effective information integration and reasoning across resources based on more relevant information closer to historical/research information.

References

[CCO05] *Cataloguing Cultural Objects (CCO): A Guide to Describing Cultural Works and their Images*, 2005. http://www.vraweb.org/CCOweb.

[DIL*06] DOERR M., IORIZZO D., LEWIS P., MARTINEZ K., ADDIS M.: Proposal for Multimedia Ontologies: ISO/FDIS 21127. In *Multimedia Ontology meeting during EWIMT* (2005).

[Doe03] DOERR M.: The CIDOC CRM – An Ontological Approach to Semantic Interoperability of Metadata. *AI Magazine 24* (2003).

[DPK*04] DOERR M., PLEXOUSAKIS D., KOPAKA K., BEKIARI C.: Supporting Chronological Reasoning in Archaeology. In *Computer Applications and Quantitative Methods in Archaeology Conference, CAA2004* (Prato, Italy, 13-17 April 2004).

[Gua98] GUARINO N.: Formal ontology in information systems. In *Formal Ontology in Information Sys-*

tems (Trento, Italy, 6-8 June 1998.), N.Guarino, (Ed.), IOS Press, pp. 3–15.

[JLT*05] JEWELL M. O., LAWRENCE K. F., TUFFIELD M. M., PRUGEL-BENETT A., MILLARD D. E., NIXON M. S., SCHRAEFELL M. C., SHADBOLT N.: Ontomedia: An Ontology for the Representation of Heterogeneous Media. In *Proceedings of Multimedia Information Retrieval Workshop 2005(MMIR 2005)* (Brazil, 2005).

[SAC*06] SINCLAIR P., ADDIS M., CHOI F., DOERR M., LEWIS P., MARTINEZ K.: The Use of CRM Core in Multimedia Annotation. In *First International Workshop on Semantic Web Annotations for Multimedia (SWAMM)* (2006).

[Tan76] TANCOCK J L.: *The Sculpture of Auguste Rodin.* Philadelphia: Philadelphia: Museum of Art and David R.Godine. 1976.

[VRA02] : *Visual Resources Association Data Standards Committee, VRA Core Categories,Version 3.0.*, 2002. http://www.vraweb.org/vracore3.htm.

[WKL*98] WEIBEL S., KUNZE J., LAGOZE C., WOLF M.: *Dublin Core Metadata for Resource Discovery*. Tech. rep., (RFC 2413, IETF), September 1998.

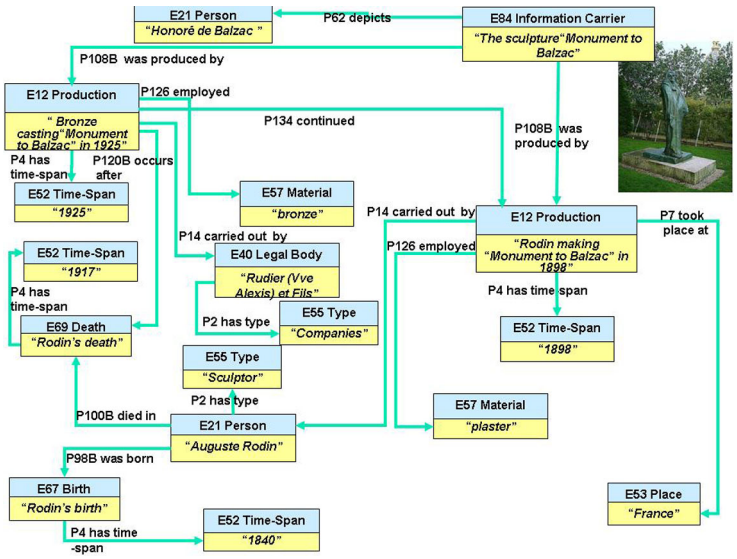

Figure 3: *A graphical representation of "Monument to Balzac" using CIDOC CRM Core*

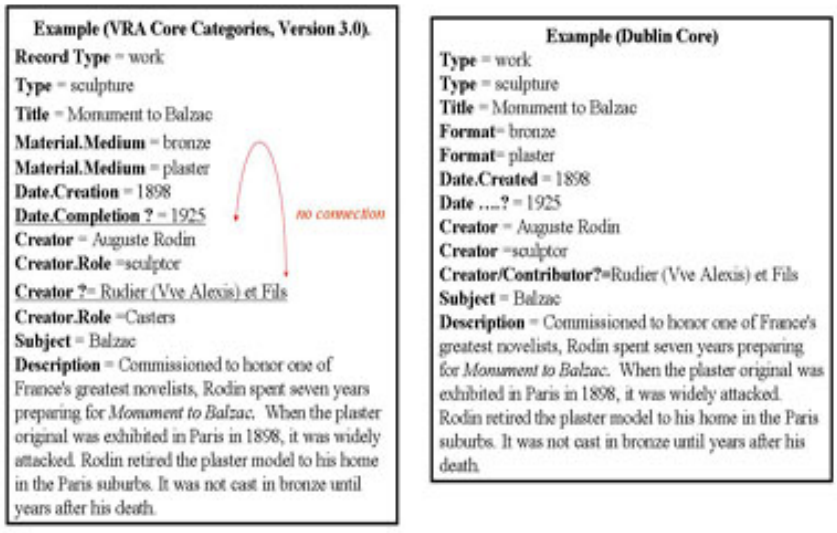

Figure 4: *"Monument to Balzac" implementing VRA and Dublin Core*