

Multi-version representation of historical event

Fairouz Zendaoui, Walid Khaled Hidouci

Laboratoire de la Communication dans les Systèmes Informatiques, Ecole nationale Supérieure d'Informatique, Algérie

Article Info

Received Dec 31th, 2018

Keyword:

History

Event

Knowledge modelling

Multi-version representation

Class diagram

ABSTRACT

The digital humanities are a field of research, teaching and engineering at the crossroads of computer science and the arts, literature, human and social sciences. Historical disciplines focus on digital tools, especially databases. Current interests and efforts focus on the representation of historical knowledge in order to facilitate the diffusion, sharing and exploitation of collective knowledge. In this context, we propose in this article a representation model of historical event. The particularity of our model is to represent simultaneously multiple versions of the same event from different sources. Furthermore, we discuss our model as a field of expertise for new research and investigation problems.

Corresponding Author:

Fairouz Zendaoui,

Laboratoire de la Communication dans les Systèmes Informatiques,

Ecole nationale Supérieure d'Informatique,

BP 68M, 16309, Oued-Smar, Alger, Algérie.

<http://www.esi.dz>

f_zendaoui@esi.dz

1. Introduction

The capitalization of history is a priori the responsibility of historians. However, it goes far beyond the historical problem when it is part of a perspective of knowledge representation. Replaced in this context, it can then be the subject of specific approaches, including through knowledge engineering. Nevertheless, this faces for the moment two types of difficulties. On the one hand, history techniques have few tools for representing historical knowledge, and are even rather reluctant to use such tools. On the other hand, knowledge engineering has not posed the problem of representing historical knowledge, tracing the evolution of history. However, it is possible to develop robust and validated methods, tools and techniques that take into account these two approaches, which, if they work in synergy, prove to be rich and fruitful.

We can obtain different versions of the same event from several testimonies more generally sources (such as archives, documents, vestiges, etc.). However, to our knowledge none of the works that have represented historical event has considered its multi-version aspect and has been content to keep only one version of the event history, the one deduced by the historian, by applying the historical process, as a definitive version.

In this paper, we propose our model of representation of historical event. Our interest is in the simultaneous preservation of multiple versions of an event from different data sources. Our goal is to make available to the historian community a ground of reflection and expertise allowing it to answer some questions in historical research.

This article is structured as follows. In Section 2, we introduce the concept of history. Then, we approach the knowledge modelling in history, and expose the related technologies and works. We articulate our contribution in section 3, where we expose our representation model of historical event and discuss some interesting uses. We conclude our paper in section 4 and state some perspectives in section 5.

2. History and knowledge modelling

2.1. Concepts and definitions in history

History is the story's daughter. It is not defined by an object of study, but by a type of speech. To say that she studies time has no other meaning than to say that she has all the objects she studies in time: to make history is to tell a story [1].

The event is about what happens at a certain date and place. It is therefore spatio-temporal. Moreover, it is what distinguishes itself from the everyday, from the banal, from the ordinary. As such, it is a fact but differs from other facts of the same nature. Then, an event is what is true for a group, a people, a culture, a large group of men. It has a proper meaning that is not only mechanical, like the facts. It is also, what modifies in the long term and irreversibly a certain human reality [2]. An event can be as much political, scientific, natural as historical.

The role of the historian is to explain historical events, to establish their relations, to define their causes and their consequences, in short to reveal the coherence, the organization of a historical period, its continuity or its discontinuity. It thus determines a trend, a logic in the apparent chaos of events. The historian must therefore identify only the historical elements that make sense to answer a previously established problem and thus indicate the meaning of history [3].

2.2. Knowledge Modelling in history

Knowledge representation involves modelling the known knowledge of experts and bringing them together in the knowledge base. This is done by identifying granules of knowledge, abstract entities containing a certain amount of knowledge and which will be implemented in adequate data structures. The representation of knowledge can thus be defined as the set of techniques that allow the manipulation of knowledge fragments [4].

The need for the historian either as a researcher or as a teacher, in the face of complex processes or systems involving a large amount of data, to draw out simplified and serial representations, has pushed the structuralist school to think about the concept of knowledge modelling according to three different methods [5].

a) The first is based on a systemic type of modelling: this is the example of R. Souriac. According to the principles of systemic modelling [6], the method consists in describing on a model the interacting subsystems, the positioning of actors, the key functions to identify the constraints allowing to explain a historical process (acting games).

b) The second is called semiotic [7]. It is based on the elementary units of meaning (“semes”, hence “semiotic”), some combinations of which are capable of providing relevant explanations of complex systems. Therefore, the method consists in identifying significant elements in the objective of the study of “signs” or “criteria” and determining a relevant combinatorial of these signs to develop intelligible models.

c) The last, in addition to the two previous ones, aims to include space in historical modelling, called historical geography or geo-history [8]. The purpose of this method is to propose an interpretation of reality, to illuminate the maps of the past and to contribute to the understanding of the transition from one map to another.

2.3. Technologies and works in modelling historical knowledge

Several researchers have represented historical knowledge to support the production and sharing of knowledge in history. The first works of historical knowledge representation have used databases with two types of interest: prosopographic databases (prosopography is the description of people based on the same model) and social networks [9]. In the majority of recent works, interest and efforts have focused on the semantic representation of history especially with the emergence of semantic web that shook the paradigm of publishing and sharing research data.

Reference [10] presents a very rich art state of the various works dealing with the use of methods and technologies of the Semantic Web in historical research. Among these works, we mention the project SYGMOGIH a collaborative platform based on a geographic information system for geo-historical resource sharing [11]. Another project called SEGRADA still based on semantic networks represents a historical crossroads exploiting a semantic database allowing the preservation of the information and the semantic links that connect them [9]. CIDOC CRM is a reference semantic model constituting an information ontology

relating to cultural heritage defined as an international standard by ISO since 2006 [12]. This standard was applied in [13] in order to develop event ontology for the First World War. Reference [14] has structured data from Thailand's historical events in an effort to improve learning in an innovative way.

The works cited did not focus on the task of multi-version representation of historical events. Either they have simply represented a single version of a historical event that is deduced by applying a historical process, or they have allowed multiple representations of the same event without linking them, which does not allow for reflect the same event or mark the multi-version aspect in a multi-source environment.

3. Representing multiple versions of historical event

In this section, we present the objectives of our modelling. Then we expose our static analysis of the model that we illustrate with a class diagram. In the end, we discuss some possible extensions and interesting uses of our model.

3.1. Principles of the model

The representation model should make it possible to formally describe the historical events so that the history of events can be reconstructed effectively. The principle of our knowledge model is to represent knowledge about the historical events by keeping multiple versions of the same historical event provided from different sources. For this, for each knowledge unit is associated the reference source and date.

We have opted for the structuring of historical knowledge by using a relational database to benefit from the numerous mechanisms allowed by database management systems, namely: ensuring the coherence of information, avoiding loss of information due to breakdowns, ensure confidentiality and allow shared use. We focused our modelling on the event. First, we identify the entities necessary for the representation of the historical event and represent them by object classes. Then, we identify the associations that link the object classes in order to reconstruct knowledge about the historical event.

3.2. Static analysis

A. Object classes identification: For covering the historical event, we deemed necessary to represent the following entities: event, place, date, actor, object, source, appellation. Fig. 1 represents a UML class diagram which only illustrates the entities of the model proposed by a generalization / specification relation.

B. Identification of associations: To link object classes and reconstruct historical event we have respected the following integrity constraints:

- An event involves one or more places.
- An event is affected by one or more temporal entities.
- An event is concerned by one or more actors.

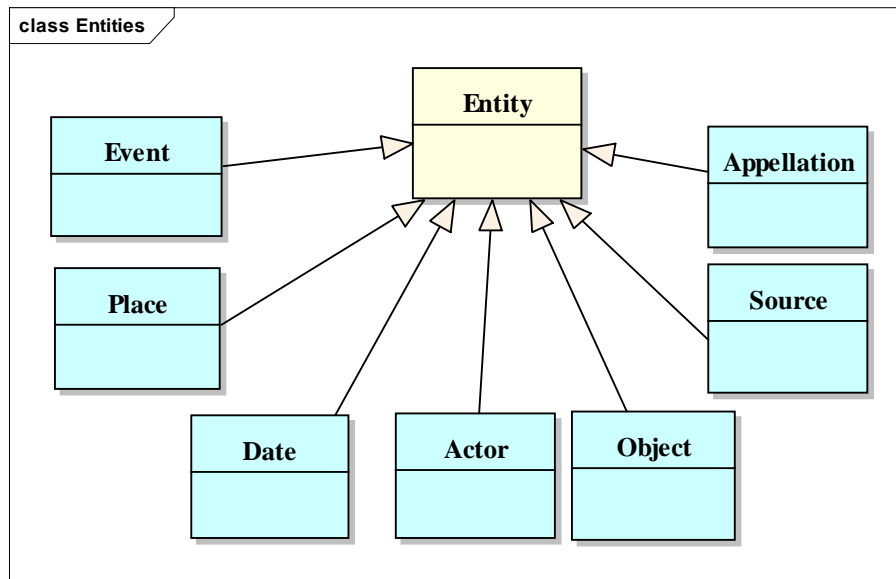


Figure 1. Class diagram representing the model entities.

- An event involves one or more objects.
- An event has one or more appellations.
- Each unit of knowledge is associated with the source and the date of reference.

Fig. 2 represents a simplified UML class diagram centred on event by abstracting other aspects such as relations between events, between places, between actors, and between sources. This diagram illustrates the relations that link each time the event to another entity by always involving the source and date of reference.

To avoid confusion, since the association that specifies the event date links the event class and the date class using two links representing the event date of the source date, we have specified the role of each link: ReferenceDate or EventDate.

Suppose that a source S1 on date SD1 asserted that event E occurred at place P1 on date D and the intervening actors are A1 and A2. Another source S2 dated SD2 stated that event E occurred on date D at place P2 and intervening actors are A2 and A3. These two sources tell different versions of the event E. They agree only for the date D of the event and the involvement of the actor A2. We will illustrate the application of our model to represent the event according to each of the two sources.

The records representing the statement of source S1 are:

- Date of the event: (E, D, S1, SD1).
- Event Location: (E, P1, S1, SD1).
- The actors of the event: (E, A1, S1, SD1), (E, A2, S1, SD1).

The registrations representing the statement of source S2 are:

- Date of the event: (E, D, S2, SD2).
- Location of the event: (E, P2, S2, SD2).
- The actors of the event: (E, A2, S2, SD2), (E, A3, S2, SD2).

C. Extensions of the model: We extend the proposed model to enrich it and bring it more semantics by modelling the following aspects:

- **Entity appellation:** Each entity has one or more appellations each characterized by the source and the date of reference. For example: a place could have a given name for a given time and that name could change over time. The trace of all these changes by the reference date and source should be conserved. The modelling of the appellation of the other entities of the model is done in the same way as we modelled the appellation of the event in Fig. 2.

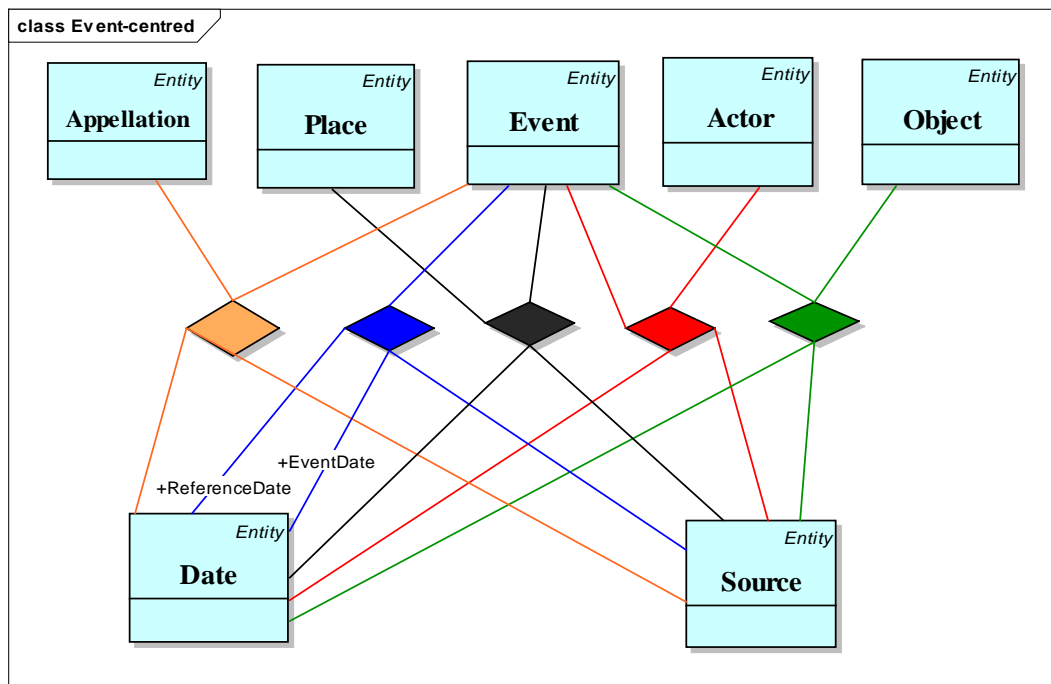


Figure 2. Event-centred class diagram.

- **Relations between events:** An event is linked to one or more other events by a relation characterized by date and source of reference (see Fig. 3). As an example of relations, we cite: hierarchy (contains / content), succession (previous / next), and causality (cause / effect).
- **Relations between places:** A place is linked to one or more other places by a relation characterized by the source and the date of reference (see Fig. 4).
- **Relations between actors:** An actor is linked to one or more actors by a relation characterized by the source and the date of reference (see Fig. 5). For example: two people could have “cousin” as a unique relation at a given time, and then add the “spouse” relation.
- **Relations between sources:** A source is linked to one or more other sources by a relation characterized by the source and the date of reference (see Fig. 6).

3.3. Exploitations and uses of the model

Our model represents a collective memory that can serve the historian community by providing as a field of expertise for new research and investigation problems:

- Detection of similarities and concordances in the different sources’ versions about a historical event.
- Detection of the source contradictions that have cited differently an event in different contexts.
- Detection of spelling conflicts related to the names of events or the others entities of the model.
- Detection of convergences and divergences of versions from different sources in relation to a historical event.

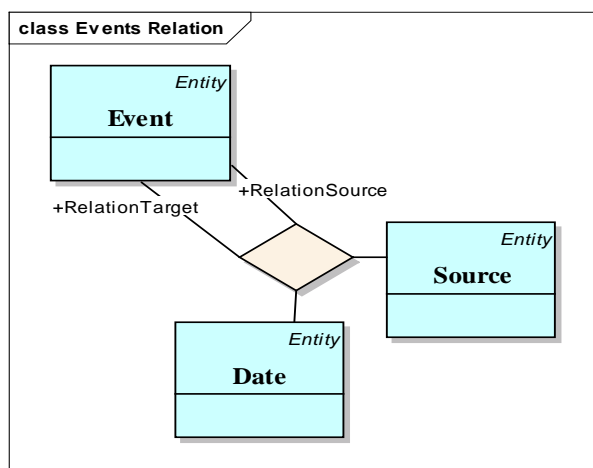


Figure 3. Events relation class diagram.

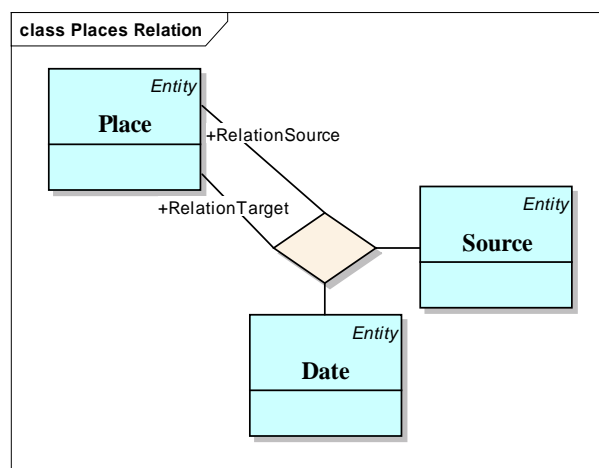


Figure 4. Places relation class diagram.

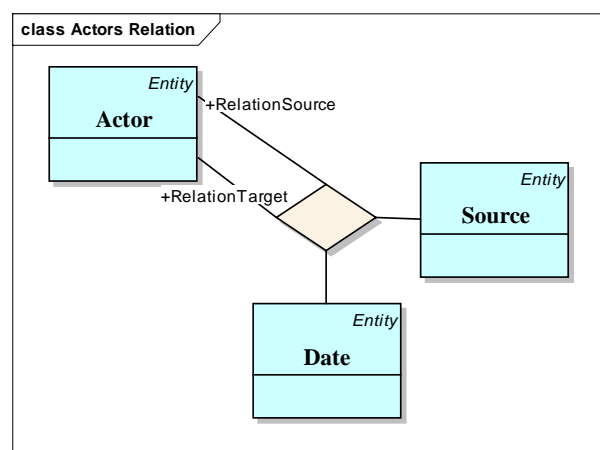


Figure 5. Actors relation class diagram.

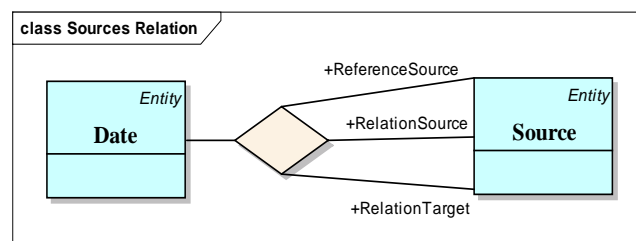


Figure 6. Sources relation class diagram.

1. Conclusion

In this article, we modelled the historical event by keeping multiple versions provided by different sources. The principle of our model is to associate to each knowledge unit the reference source and date. First, we have identified the necessary entities to cover the historical event and represented them by object classes. Then, we have identified the associations that link the object classes in order to reconstruct knowledge about the historical event. Further, we proposed some extensions of the model to enrich it and bring more semantics. We also brought out several interesting exploitations and uses of our model.

2. Perspectives

In future work, we will firstly consider the quality of imperfection of the historical data in terms of uncertainty. We will try to represent the historical event by keeping multiple versions from different sources and by associating with these versions degrees of belief corresponding to the degrees of certainty expressed by the sources. On the other hand, we will try to join a domain ontology such as CIDOC CRM to strengthen, formalize and standardize the semantic aspect of the model.

References (11 PT)

- [1] Furet F. “L’Atelier de l’historien”, Collection Champs, Flammarion, Paris, 1982.
- [2] Granda S. “Critique et Analyse des Bases de Données sur les Catastrophes Naturelle pour la Caractérisation des Evènements Hydro-Météorologique a Forts Impacts”, 2014.

- [3] Ollier-Pochart E. “Les Ecritures de L’histoire dans les Romans Québécois de la Décennie 1980-1990”, 2012.
- [4] Jouve C. “Représentation des Connaissances pour les Problèmes de Conception. Application à un Système à Base de Connaissances pour la Conception de Réseaux Informatiques: NEST”, 1992.
- [5] Ermine J. L., Pauget B., Beretti A., Tortorici G. “Histoire et Ingénierie des Connaissances. Sources et Ressources pour les Sciences Sociales”, Dec 2004, Paris, France. pp.1-15, 2004. <hal-00432738>
- [6] Le Moigne J.-L. “La Modélisation des Systèmes Complexes”, Afcet Systèmes, Dunod, Paris, 1990.
- [7] Soulet, J. F. “L’Histoire Immédiate”, Paris, PUF, coll. « Que sais-je ? », n°2841, 1994.
- [8] Grataloup C. “Lieux d’Histoire, Essai de Géohistoire Systématique”, Gip Reclus, 1996.
- [9] Kalus M. “Semantic Networks and Historical Knowledge Management: Introducing New Methods of Computer-based Research”. vol. 10, no. 3, December 2007.
- [10] Meroño-Peñuela A., Ashkpour A., Erp M., Mandemakers K., Breure L., Scharnhorst A., Schlobach S., and Harmelen F. “Semantic Technologies for Historical Research: A Survey”. *Semantic Web (2014)* 1–27, 2014.
- [11] Beretta F., and Butez C.C. “The SyMoGIH project and Geo-Larhra: A Method and a Collaborative Platform for a Digital Historical Atlas”. *Digital Humanities 2014*, Jul 2014, Lausanne, Switzerland. EPFL, Lausanne / UNIL, Lausanne, pp.117-119, 2014, *Digital Humanities*, 2014.
- [12] Doerr M. “The CIDOC CRM, an Ontological Approach to Schema Heterogeneity.” *Semantic Interoperability and Integration*, 2005.
- [13] Hyvönen E., Lindquist T., Törnroos J., and Mäkelä E. “History on the Semantic Web as Linked Data - An Event Gazetteer and Timeline for World War I”, 2012.
- [14] Ekkachai S.J., and Kulhida T. “Development of Semantic Web for Thai Historical Events”. *Proceedings of the Doctoral Consortium at the 18th International Conference on Asia-Pacific Digital Libraries, and Asia-Pacific Forum of Information Schools*, 2016.