

anthology, or should both be an integral version? Should one of the two media be favoured? Should strict philological requirements be applied to both versions?

- **establishing who should be the favoured website-user:** a pleasure-reader, a scholar looking for information, a teacher in search of educational material, or all three, thanks to a visualisation of the text chosen depending on what each visitor is aiming at?
- **combining the requirements of durability and the multiplicity of uses:** the necessity of freezing and conserving a text as a patrimony is opposed to its necessity to adapt to different kinds of distribution and reading (Microsoft Word version, “e-book” version, extraction *via* a search engine, exportation to other textual analysis tools, etc.)
- **creating a reading interface adapted to a specific text,** but compatible as much with traditional philological procedures as with the explicit and implicit criteria of Internet navigating and visualising and indexing the pages produced, while complying with a universal standard, TEI.

The Exhibition Problem. A Real Life Example with a Suggested Solution

Øyvind EIDE

*Unit for Digital Documentation at the Faculty
of Arts, University of Oslo*

Background

In a paper presented at ACH/ALLC 2005, Allen H. Renear et.al. describe a problem of potentially great significance (Renear 2005). They argue that:

“In ordinary linguistic communication we often use a name to refer to something in order to then go on to attribute some property to that thing. However when we do this we do not naturally construe our linguistic behavior as being at the same time an assertion that the thing in question has that name. (Ibid, p. 176)”

Further, they claim that this distinction is over-looked when conceptual models based on encoded texts are developed.

In our work at the Unit for Digital Documentation at the University of Oslo, we have used XML encoded material as sources for several of our databases (Holmen 1996, Holmen forthcoming). The way this is done is by marking up texts both descriptively and interpretatively, followed by the use of software to extract information which is included in the databases. If Renear’s argument is correct, we may infer that the databases include assertions which are based on information in the source texts that is, strictly speaking, not grounded in these texts. For example, we could be using a text as the source of a naming in the database while the naming is merely exhibited, and not asserted, in the text.

The false resolutions

Renear et.al. propose three possible resolutions to this problem, but they also state that all of these are false. Their resolutions are the following:

1. TEI encoding represents features of the text only.

2. The use of two arcs, i.e. “The Semantic Web community solution”, which will be discussed below.
3. Exhibition is a special case of presupposition.

Based on the description of our work above, it should be obvious that resolution no. 1 is not an alternative for us. Semantic modelling of the real world on the basis of descriptions in texts is part of our work.

I find it difficult to understand how resolution no. 3 may represent a possible solution. Whether exhibition is a type of presupposition or not does not change the basic problem; i.e. in our case, the use of a text as the source of a naming which is merely exhibited in the text. The problem remains the same if the naming is also presupposed in the text, as long as it is not asserted.

I claim that resolution no. 2 is not false after all, and below I will demonstrate how the Conceptual Reference Model (CIDOC-CRM) will solve a similar problem in my example text. The CIDOC-CRM is a ontology developed in the museum community to be used for cultural heritage documentation.

My example text

In this paper, no general solution to the problems identified above will be proposed. However, I believe that the special solution that I propose could easily be generalized.

The text used in my example is based on the work of Major Peter Schnitler. In the 1740s, Major Schnitler was appointed by the Danish government to explore the border area between the northern parts of Norway and Sweden/Finland. Significant parts of the text in the manuscript that he handed over to the Danish government consist of transcripts of local court interviews which were carried out by Schnitler in order to gather information about the local population as well as what they had to say about the border areas. The material includes information directly relevant to the border question, as well as general information of the areas in question, which corresponds to similar material collected through work carried out in Europe at the time (Burke 2000, pp. 128 f.).

The text fragments below are taken from the very first meeting described in the text (English translation from Danish by me):

- [1] Of the Witnesses, supposed to be the most Cunning

on the border issue, Were and stood up in the court
1: Ole Larsen *Riise*.

- [2] For these the Kingly order was read out loud [...] and they gave their Bodily Oath
- [3] Question: 1: What his name is? *Answer: Ole Larsen Riise* (Schnitler 1962, p. 1)

In these quotes, we find that several facts are asserted by the text.

Excerpt 1 claims the existence of a witness. We will call this witness x. Being a witness implies being a person. Thus, x is a person. We may also note that x is referred to by using the name “Ole Larsen *Riise*.”, abbreviated “OLR” below.

Excerpt 2: Person x gave an oath to speak the truth.

Excerpt 3: Person x, according to the text, claims that his name is OLR. The source of the naming is person x, as spoken out loud at a specified place at a specified date in 1742. The text puts forward an assertion by person x that he is named OLR.

Modelling the semantic content from our perspective

My semantic model of these facts will include the following information:

Assertion	Source
1) There is an x who is a witness	The text
2) x is a person	The meaning of the word “witness” and “person” in this context
3) x gave an oath	The text
4) OLR is the name of person x	x

It is easy to describe the source of the three first assertions through CIDOC-CRM, by stating that they are documented in Schnitler’s text:

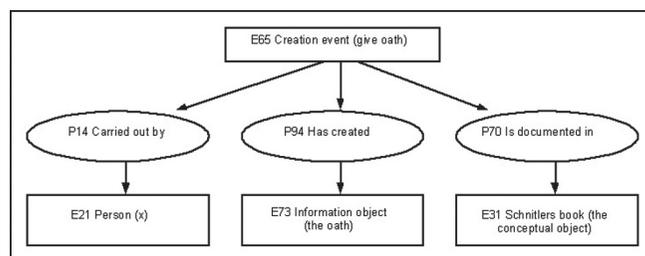


Figure 1

In this figure, as well as in the next one, the boxes with names starting with *E* represents entities, while the boxes with names starting with *P* represents the properties linking them together.

But how do we describe the source of the naming event? We start with the event in which the attribute was assigned (the naming event, a speech act), which is an *E13 Attribute assignment* which states that *x* carried out this particular speech act:

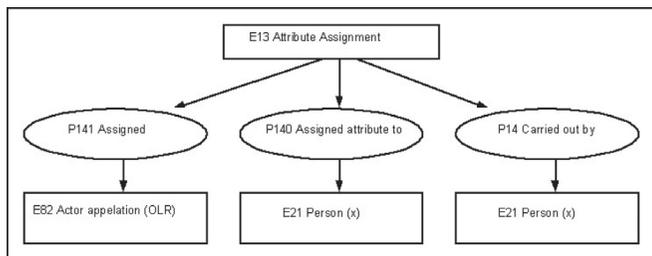


Figure 2

When looking at these two model figures, it is striking to what extent the modelling of the giving of the oath in Figure 1 compares to the naming of *x* in Figure 2. The explanation is that those are similar situations. Our traditional way of reading made us structure them differently in the table above, whereas represented in the CIDOC-CRM structure they came out the same in Figure 1 and 2. In order to show clearly in what way they correspond, note that line 4 in the table above could be rewritten as follows:

Assertion	Source
4) <i>x</i> named himself ORL	The text

This is a good example of the way modelling may help us understand a text better. What we have done is to rethink the difference between an event (*x* gave an oath) and a fact (ORL is the name of *x*). In order to model the fact correctly, i.e. to show that it was exhibited rather than asserted in the text, we had to consider it as a naming event. Considering it as an event is more feasible in that an event typically has actors who are responsible for the outcome. Further, this makes more sense in that both expressions are speech acts. When it is considered as a speech act, the naming event is the same kind of

event as the giving of an oath.

Why solution 2 is not false after all

In order to be able to see the problem with Renear’s solution no. 2, or to realize that the problem is not really there, we have to quote his text in extenso:

“Another approach, this one anticipated from the Semantic Web community, is simply to insist on an unambiguous corrected conceptual representation: one arc for being named “Herman Melville”, one for authoring *Moby Dick*. But this resolution fails for the reasons presented in the preceding section. Although this model would be in some sense an accurate representation of “how the world is” according to the document, it would not represent what is asserted by the document. The authorship arc in the corrected RDF graph model will correspond to relationships of exhibition, not assertion; and there is no accommodation for this distinction in the modelling language. (Renear, p. 178)”

In the first couple of sentences in this paragraph, the resolution of using an “unambiguous corrected conceptual representation” is said to have failed. The next couple of sentences weakens his statement by saying that only RDF does not accommodate this; “there is no accommodation for this distinction in *the* modelling language” (my emphasis). There are no arguments to support why a different modelling language could not solve the problem. In fact, the CIDOC-CRM does solve this, by giving the modeller an opportunity to state explicitly who is the source of an assertion, as demonstrated in Figure 2.

In the example above, we knew who made the assertion exhibited in the text. But even if we did not know, we could still make a similar model as long as we accept that it was made by somebody. In CIDOC-CRM, the modelling of entities we infer to exist without knowing who or what they are is quite possible.

Generalization

The example described above is quite special, as it includes an explicit naming. But it can be argued that all person names, at least in 18th century Scandinavia, are based on naming events, as people are baptised. As long as we believe that this is the case, we can include

in the model an explicit attribute assignment event as the one in Figure 2 for each name used in the text. This will be an event of which we do not know who carried it out or when it took place, but that is not necessarily a problem. There will always be things we do not know in historical texts. The naming event we model this way will also be an event that is not documented in the text we are basing the model on. Whether this is acceptable is a decision one has to take when building up such a model.

Conclusion

There is reason to believe that the problem described in Renear's paper is an important one. But a solution to the problem has been identified. I have shown that for one specific type of text, the problem may be solved by using CIDOC-CRM modelling including explicit statements of the sources of the assertions exhibited in the text. Further research may disclose whether this solution will work for other types of texts as well.

References

- Burke, P.** (2000) *A social history of knowledge : from Gutenberg to Diderot*. Cambridge.
- CIDOC-CRM.** *ISO/FDIS 21127. Information and documentation -- A reference ontology for the interchange of cultural heritage information* [Definition of the CIDOC Conceptual Reference Model].
- Holmen, J.; Uleberg, E.** (1996) "Getting the most out of it - SGML-encoding of archaeological texts." Paper at the IAAC'96 Iasi, Romania. URL: http://www.dokpro.uio.no/engelsk/text/getting_most_out_of_it.html (as of 2005-11-14).
- Holmen, J.; Jordal, E.K.A; Olsen, S.A.; Ore, C.E.** (forthcoming) "From XML encoded text to objects and events in a CRM compatible database. A case study". In: *Beyond the Artifact. Proceedings of CAA 2004, Computer Applications and Quantitative Methods in Archaeology*.
- Parsons, T.** (1990) *Events in the semantics of English : a study in subatomic semantics*. Cambridge, Mass.
- Renear, A.H.; Lee, J.H.; Choi, Y.; Xiang, X.** (2005) "Exhibition: A Problem for Conceptual Modeling in the Humanities". P. 176-179 in: *ACH / ALLC 2005. Conference Abstracts*. 2nd Edition, Victoria.
- Schnitler, P.** (1962) *Major Peter Schnitlers grenseeksaminasjonsprotokoller 1742-1745. Bind 1 [Major Peter Schnitlers border examination protocols 1742-45]* / by Kristian Nissen and Ingolf Kvamen. Oslo.