tangible, as well as the relations types that have been defined. Thus we can unveil the social networks that exist in a hidden state in the documents. The visualisation in the form of a graph proves the usefulness of ontologies when encoding an index nominum. Because a relational database cannot show relations other than hierarchical, or deduce inferences of expressed relations, it cannot completely realise the modelling of social networks. On the contrary, an ontology allows this process and transforms the index from a classical access point into an real tool for analysing the corpus. The social networks revealed by the ontology in the graph form become easier to apprehend, and this will constitute, once this technology is implemented on significant historical corpus, an essential manner to study the presence of a group of people in a geographic space according to their relations. The setting up of the index nominum as an ontology relying on XML structured works shows how a technology like RDF can grow the interest of a well identified tool using the modelling of information.

Axiomatizing FRBR: An Exercise in the Formal Ontology of Cultural Objects

Allen RENEAR

Yunseon CHOI

Jin Ha LEE

Sara SCHMIDT

Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign

Most current conceptual modeling methods were originally designed to support the development of business-oriented database systems and cannot easily make computationally available many of the features distinctive to cultural objects. Other modeling approaches, such as traditional *conceptual analysis* can complement and extend contemporary conceptual modeling and provide the computing humanist with methods more appropriate for cultural material and humanistic inquiry.

The Humanities and the Problem of Method

Dilthey famously distinguished the methods of the cultural sciences from those of the natural sciences, claiming that the natural sciences seek to explain whereas the sciences of culture seek to understand as well. Although there is no generally accepted account of this distinction, it is still a not uncommon belief that when humanists analyze, explain, and interpret the cultural world, they are, at least in part, using distinctive methods. The question has a long history but it is now especially acute in the practice of humanities computing.

One compromise is to accept the separation and treat computational support as preliminary or ancillary — or, even if constitutive, partial, and the lesser part. We believe that such a resolution will result in missed opportunities to develop intrinsic connections between the methods of managing computational support and traditional methods of advancing humanistic insight.

Conceptual Analysis

The early Socratic dialogues focus on cultural concepts such as justice, piety, courage, beauty, friendship, knowledge, and so on. Socrates asks what these are and attempts to determine what features are significant, sometimes considering hypothetical cases to elicit modal intuitions, sometimes reasoning discursively from general principles. This now familiar style of reasoning may be called "conceptual analysis", or, when formalized with the general principles articulated as axioms, "axiomatic conceptual analysis". Often discussions of cultural objects by humanist scholars can be seen to be some variation of this sort of reasoning or situated within a framework of concepts which could be explicated in this way.

This approach to understanding cultural facts has been widely criticized, from both hermeneutic and positivist quarters; however recent work on the nature of social facts may provide some support. Searle and others have argued that social and cultural facts are established through acts of "collective intentionality" (Searle, 1995). If so then at least part of the nature of that reality would seem to be directly accessible to the participating agents. We cannot investigate galaxies, electrons in this way, because we in no way create them as we do poetry, music, and social institutions. Searle's account is consistent with the approach taken by the phenomenologists of society and culture, such as Reinach and Ingarden, as well as with classical Anglo-American philosophical analysis (Smith, 2003).

An Example: Bibliographic Entities

In the *Functional Requirements for Bibliographic Records* (FRBR) the conceptual modeling is explicit and the conceptual analysis latent. In the text of FRBR we read (IFLA, 1998):

Work: "a distinct intellectual or artistic creation"

Expression: "the intellectual or artistic realization of a work in the form of alphanumeric, musical, or choreographic notation, sound, image, object, movement..." (e.g., a text).

Manifestation: "the physical embodiment of an expression of a work". (e.g., an edition).

Item: "a single exemplar of a manifestation". (e.g., an individual copy of a book)

The novel *Moby Dick*, a work, is realized through various expressions, the different texts of Moby Dick, including different translations. Each one of these expressions may be embodied in a number of different manifestations, such as different editions with different typography. And each of these manifestations in turn may be exemplified in a number of different items, the various individual copies of that edition. Each entity is also assigned a distinctive set of attributes: works have such things as subject and genre; expressions a particular language; manifestations have typeface and type size; and items have condition and location.

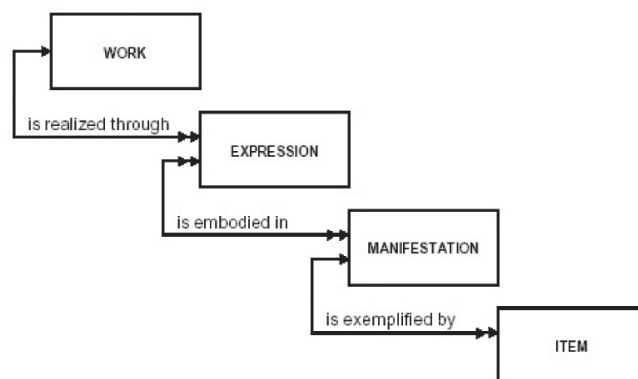Below is the "entity relationship diagram" representing these entities and relationships:

Figure 1: ER Diagram of FRBR Group 1 Entities and Primary Relationships
[diagram from IFLA (1998)]

Entity relationship diagrams are a widely used conceptual modeling technique in the development of information management systems and there are algorithms for converting ER diagrams into robust lower level abstractions, such as normalized relational tables, that can be implemented in database systems. However standard ER diagrams cannot make all aspects of cultural material computationally available. There is no method for saying explicitly under what formal conditions entities are assigned to one entity set or another, for distinguishing entities from relations and attributes, or for identifying necessary or constituent features. Moreover, relationships are understood extensionally, and modal or other intentional assertions, including propositional attitudes and speech acts that are critically important in

the study of society and culture cannot be expressed. (Renear and Choi, 2005).

Extending Conceptual Modeling with Conceptual Analysis

[*Caveat:* In what follows we intend no position on the plausibility of any ontological theory of cultural objects. Our claim is only that there are such positions, that they cannot be easily represented with current conceptual modeling techniques, and that they can be represented with other techniques.]

The text of FRBR provides much information that, despite appearances, is not represented in the FRBR ER diagram. Some of this disparity has been discussed elsewhere (Renear and Choi, forthcoming); here we take up features especially relevant to cultural material.

For example, the FRBR ER diagram does show embodiment, realization, and exemplification relationships, of course, but it does not indicate their particular significance. The "is" of "...is the physical embodiment..." is not the "is" of mere predication. It is a conceptually constitutive "is": we are being told not just a fact about manifestations, but what manifestations (conceptually) *are*. The cascade of definitions suggests this formalization:

work(x)   … x is an artistic or intellectual creation

expression(x)   =df ($\exists y$)[realizes(x,y) & work(y)]

manifestation(x) =df ($\exists y$)[embodies(x,y) & expression(y)]

item(x)   =df ($\exists y$)[exemplifies(x,y) & manifestation(y)]

Now we can see that the concept of work is taken as a quasi-primitive entity, the three characteristic relationships are also each primitive, and essentially involved in the definitions of the entities, and the appearance of interdefinition is made explicit. Because none of this is modeled in the FRBR ER model, that model does not fully represent FRBR's perspective and, moreover, these features will not be reflected in information systems generated from that model and will not be computationally available for analysis.

Bibliographic Platonism

Already we see inferences not entirely trivial, such as the theorem that bibliographic items imply the existence of corresponding (abstract) manifestations, expressions, and works:

P1   item(v) ⊃ ($\exists x$)($\exists y$)($\exists z$) [manifestation(x) & expression(y) & work(z) & exemplifies(v,x) & embodies(x,y) & realizes(y,x)]

But consider the converse of that conditional:

A1   work(v) ⊃ ($\exists x$)($\exists y$)($\exists z$) [item(x) & manifestation(y) & expression(z) & exemplifies(v,x) & embodies(x,y) & realizes(y,z)]

A1, a bibliographic analogue of the Aristotelian thesis that only instantiated universals exist does not follow from the definitions.

Represented in this way FRBR now raises a traditional problem for Platonist ontologies of art: if works are abstractions existing independently of their instantiations, then how can they be *created*?

Bibliographic Aristotelianism

An alternative approach could take items as primitive.

work(x)   =df ($\exists y$)[IsRealizeBy(x,y) & expression(y)]

expression(x)   =df ($\exists y$)[IsEmbodiedBy(x,y) & manifestation(y)]

manifestation(x) =df ($\exists y$)[IsExemplifiedBy(x,y) & item(y)]

item(x)   =df a (material) artistic or intellectual creation

This is a "moderate" realism in which **A1** is now a theorem and **P1** no longer one. Here abstract objects cannot exist independently of their physical instantiations, although they do exist (as real objects) when their corresponding items exist. However **P1** will certainly need to be added as an axiom to support our intuition that items do imply works in any case. Or, another approach to the same end is to leave work as primitive, as before, but add **P1** as an axiom. Either might better fits our commonsense intuitions about artistic creation, But we may now have problems characteristic of moderate realism: how to exclude abstract objects which have an intermittent being, going in and out of existence as their instances do — which would be in contradiction to another commonsense intuition: that "a thing cannot have two beginnings in time" (Locke).

A Third Way

In FRBR the notion of a work seems poorly accounted for, tempting further development. Jerrold Levinson defends this definition of musical work:

x is a musical work =df

x is a sound/performance_means-structure-as-indicated-by-S-at-t.

Levinson argues that works are "initiated types" (other examples: the Ford Thunderbird and Lincoln penny) which do not exist until *indication* but once created exist independent of their concrete instances. Our intuitions about artistic creation are now accommodated, but at a cost: a special class of abstract object which, at least arguably, has a beginning in time but never an end, as in Karl Popper's "third world" of cultural objects. Revising the formalization to represent this view is left to the reader as an exercise. It is a little harder than you might think.

International Federation of Library Associations (1998)

Levinson, J (1980). "What a Musical Work Is," *The Journal of Philosophy* 77.

Renear, A. and Choi, Y. (2005) "Trouble Ahead: Propositional Attitudes and Metadata". Proceedings of the 68th Annual Meeting of the American Society for Information Science and Technology (ASIST). Charlotte NC.

Renear, A. and Choi, Y. (forthcoming) "Modeling our Understanding, Understanding Our Models — The Case of Inheritance in FRBR".

Searle, J. (1995). *The Construction of Social Reality*. 1995.

Smith, B. (2003) '"John Searle: From Speech Acts to Social Reality", in B. Smith (ed.) *John Searle*. Cambridge University Press.

Collaborative Scholarship: Rethinking Text Editing on the Digital Platform

Massimo RIVA
Vika ZAFRIN

Brown University Italian Studies

INTRODUCTION

Based at Brown University, the Virtual Humanities Lab [1] is one of twenty-three "models of excellence" in humanities education, supported by the National Endowment for the Humanities for 2004-06. [2] The project is being developed by the Department of Italian Studies in collaboration with Brown's Scholarly Technology Group and with scholars in the U.S. and in Europe.

This paper will report on the achievements of VHL's work during the first two years of its existence as a platform for collaborative humanities research. We will discuss the editing process as we envision it: as a form of interdisciplinary and collaborative knowledge work. We will present issues arising from our experiment with subjective (or "idiosyncratic") text encoding; challenges we face in organizing the work of an international group of collaborators and the procedures for that work; and the process of annotating and indexing large texts collaboratively. Finally, we will hint at VHL's potential applications for pedagogical purposes.

TEXT ENCODING

We have made three Early Modern Italian texts available online: Giovanni Boccaccio's Esposizioni sopra la Comedia di Dante; portions of Giovanni Villani's Cronica Fiorentina; and Conclusiones Nongentae Disputandae by Giovanni Pico della Mirandola. These three texts were selected as representative of different textual typologies (commentary, chronicle and treatise) that solicit different encoding and annotating strategies. The first two are large (around 700 and 200 modern print pages respectively) and heavily semantically encoded.