Ontologies, Multi-Perspective Modelling and Knowledge Auditing

John Kingston

Artificial Intelligence Applications Institute
Division of Informatics
University of Edinburgh

Abstract. I have extended an existing ontology for knowledge acquisition to support the task of knowledge auditing. This has required two types of change: the ontology has been extended according to the principles of multi-perspective modelling so that it represents knowledge from the viewpoints of ”who”, “what”, “where” and “how”; further work may add “when” and “why” perspectives. It also adds some slots to the class of Publications, which are intended to store information for the purpose of knowledge valuation.

1 Introduction

This paper is a position paper describing work being carried out under the EPSRC-funded Advanced Knowledge Technologies project. It describes an attempt to apply the philosophy of multi-perspective modelling [1] to an existing ontology of knowledge resources, in order to extend that ontology in sensible ways. It also touches on further extension of the ontology to permit evaluation of knowledge resources. The overall aim of the work currently being carried out is to support a knowledge audit – that is, a survey of the knowledge that is available in an organisation.

The goal of this work described in this paper has been to produce an ontology that represents all the aspects of multi-perspective modelling; that is, for any knowledge resource, it can represent *what* it is, *who* possesses it, *how* it is used, *where* it can be found, *when* it is needed and *why* it exists (or why it is useful). If knowledge is collected and is indexed according to all aspects of the above ontology, then it should be possible to browse all the people who possess a particular knowledge resource (or part of it); or all the knowledge resources held by a particular person; or all the activities that can be supported by a particular knowledge resource.

However, a good knowledge audit does not just produce a survey report; it must provide estimates of the “goodness” or “value” of knowledge, and it also ought to encode a number of relationships between knowledge of different types, in order to allow browsing of knowledge from different viewpoints or different foci of interest. This paper briefly discusses the principles of multi-perspective modelling and of knowledge valuation, and then describes the ontology that has been developed, with a couple of examples of its use. The examples are drawn from the self-referential domain of AI researchers, their research and their collaborators – specifically, the researchers and industrial collaborators on the AKT project. The “semantic web researchers” ontology published by the KA2 project was used as a starting point.

2 Multi-Perspective Modelling

The thesis of multi-perspective modelling is that for any “knowledge asset” to be represented adequately, it’s necessary to represent a number of different perspectives on its knowledge – and, possibly, to represent the asset at multiple different levels of decomposition. These ideas are based on those of the Zachman framework [2], and are embodied in various knowledge modelling methods, notably the CommonKADS methodology for knowledge engineering [3].

The Zachman framework (also called the Information Systems Architecture framework) has six columns representing *who, what, how, when, where* and *why* perspectives on knowledge, and six rows representing different levels of abstraction (see Table 1). Zachman illustrates the different levels of abstraction using examples from design and construction of a building, starting from the “scope” level (which takes a “ballpark” view on the building which is primarily the concern of the architect, and may represent the gross sizing, shape, and spatial relationships as well as the mutual understanding between the architect and owner), going through the "enterprise" level (primarily the concern of the owner, representing the final building as seen by the owner, and floor plans, based on architect's drawings) and on through three other levels (the "system" level, the "technology constrained" level and the "detailed representation" level, respectively the concerns of the designer, the builder and the subcontractor) before arriving at the "functioning enterprise" level (in this example, the actual building). Zachman describes this framework as *"a simple, logical structure of descriptive representations for identifying 'models' that are the basis for designing the enterprise and for building the enterprise's systems"* [2].

So the thesis of multi-perspective modelling is that, in order to capture all the important aspects of a knowledge asset, it is necessary to model (or at least, consider modelling) each of the six columns of the Zachman framework.

Table 1. The Zachman framework (from [2])

	Data **"*what*"**	**Function** **"*how*"**	**Network** **"*where*"**	**People** **"*who*"**	**Time** **"*when*"**	**Motivation** **"*why*"**
Objectives/ Scope "*contextual*"	List of things important to the business	List of processes the business performs	List of locations in which the business operates	List of Organizations important to the business	List of Events significant to the business	List of Business goals/ strategies
Enterprise "*conceptual*"	e.g. Semantic Model	e.g. Business process Model	e.g. Business legacy systems	e.g. Work Flow model	e.g. Master Schedule	e.g. Business Plan
System "*logical*"	e.g. Logical data model	e.g. Application Architecture	e.g. Distributed Systems Architecture	e.g. Human Interface Architecture	e.g. Processing Structure	e.g. Business Rule Model
Technology constrained "*physical*"	e.g. Physical data model	e.g. System Design	e.g. System Architecture	e.g. Presentation Architecture	e.g. Control Structure	e.g. Rule design
Detailed representations "*out-of-context*"	e.g. Data description	e.g. programs	e.g. Network architecture	e.g. Security Architecture	e.g. Timing Description	e.g. Rule Specification
Functioning enterprise	e.g. Data	e.g. Function	e.g. Network	e.g. Organization	e.g. Schedule	e.g. Strategy

3 Valuation of Knowledge

Valuing knowledge is a difficult task, because measuring knowledge is a difficult task, and determining the contribution of knowledge to an activity is also difficult. Some people have taken an approach of valuing products and partial products, calculating the knowledge required to produce each product as a function of the training time and/or salaries required by practitioners, and then generating statistics such as "return on knowledge" (see e.g. [4]). Others (e.g. England's North Western AI Applications Group) have focused on *knowledge items* instead [5], but concentrated less on the content of knowledge items than on the relationships with other knowledge items – e.g. "you need to know A before you can understand B". Such a "knowledge map" could be used in conjunction with acquisition times for each knowledge item to determine approximate total acquisition times for certain knowledge items, and thus to calculate an "opportunity cost" for them.

4 Our Approach

I have devised an approach to knowledge auditing which uses an ontology of knowledge-related terms. The aim is to carry out a knowledge audit of AI research and researchers, and so the terms focus on research topics, publication details, and so on. I used the KA2 Knowledge Acquisition Community ontology [6] as a starting point.[1]

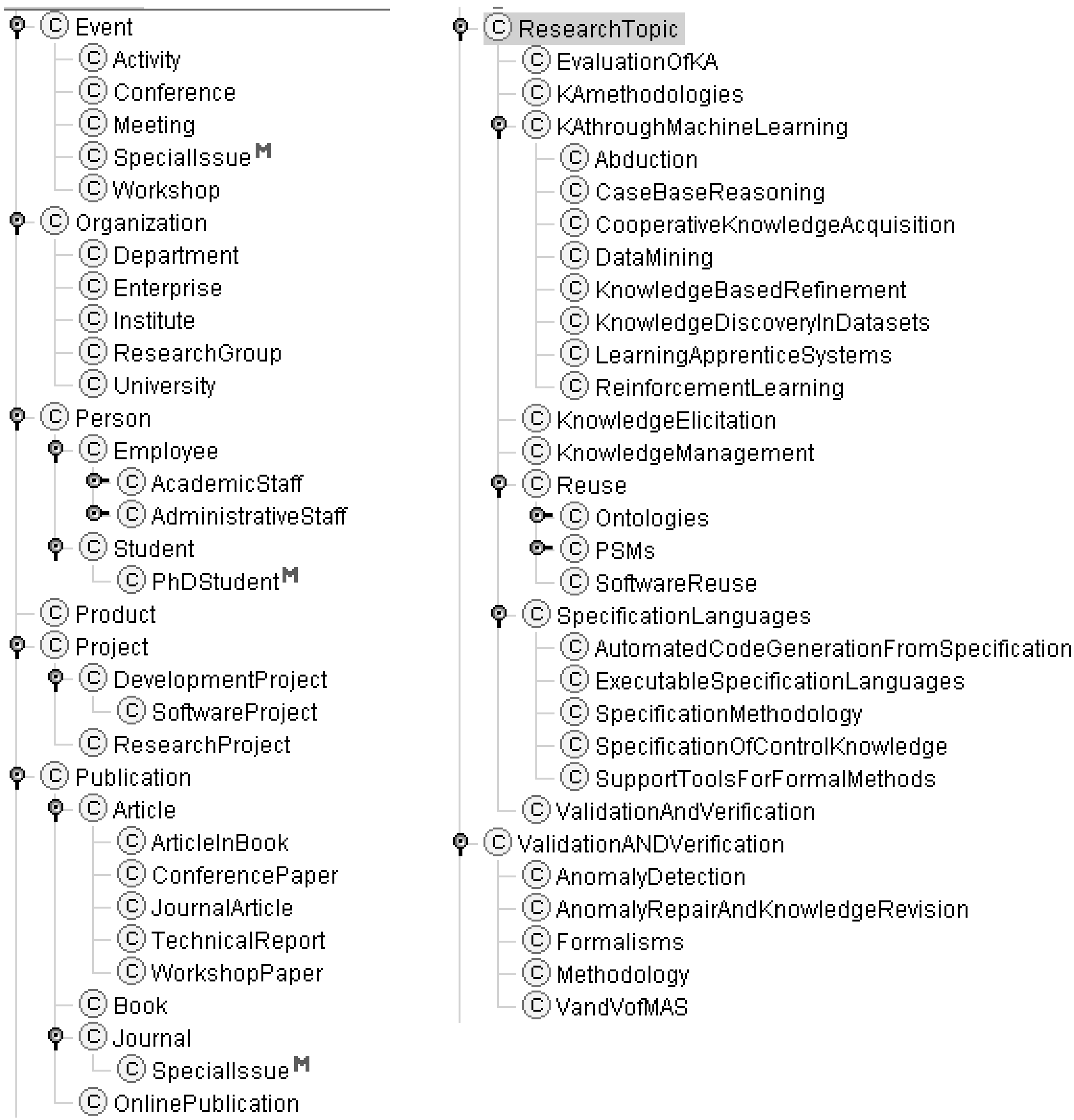

Fig. 1. Top 3 levels of the Knowledge Acquisition Community ontology

The Knowledge Acquisition Community ontology provides the following top level classes: Event, Organisation, Person, Product, Project, Publication, ResearchTopic[2]. I consider that this provides two or three of the six recommended perspectives; the ResearchTopic is the domain ("what") of a knowledge item, the Person is the owner ("who"), and the location ("where") of the knowledge is likely to be in a Publication or a Project. I have added an ontology of tasks that need to be performed, which outlines *how* certain top level tasks are carried out; for example, "developing a KBS" breaks down into

[1] Specifically, I downloaded the OIL version, and read it into Protégé-2000 as an RDF Schema file. Some additional work was necessary to recreate slot-to-class links which did not translate correctly.

[2] ValidationANDVerification is also given as a top level class. I consider this to be poor modelling and have made it a sub-class of ResearchTopic – see Figure 2.

several subtasks, which are themselves decomposed down to the level of tasks such as "detecting circularity in rule based systems". As yet, there are no specific ontologies of *when* knowledge is created or used (the date of a publication is considered sufficient) nor of *why* knowledge exists or is needed, but such ontologies could be developed if needed – for example, for time-critical applications or applications where storage space or acquisition time is limited.

Using this ontology, it should be possible to encode every relevant piece of information about a knowledge item. For example, the ontology describes the publication details of the paper "COVERAGE: Verifying Multiple-Agent Knowledge-based Systems" [7] covers the research topics of Verification and Validation, Anomaly Detection and Agent Based Systems; the task supported is Anomaly Detection in Multi Agent Systems; and the knowledge owner is the author of the paper (Alun Preece). The paper was published in a journal, whose details are duly recorded, and its publication year was 1999, so it is recent work. An ontology such as this that contained details of all the papers by a large group of researchers should provide the ability to move seamlessly between different views on knowledge; e.g. "show me all the people who know about agent based systems" or "show me all the work on KBS design after 1996".

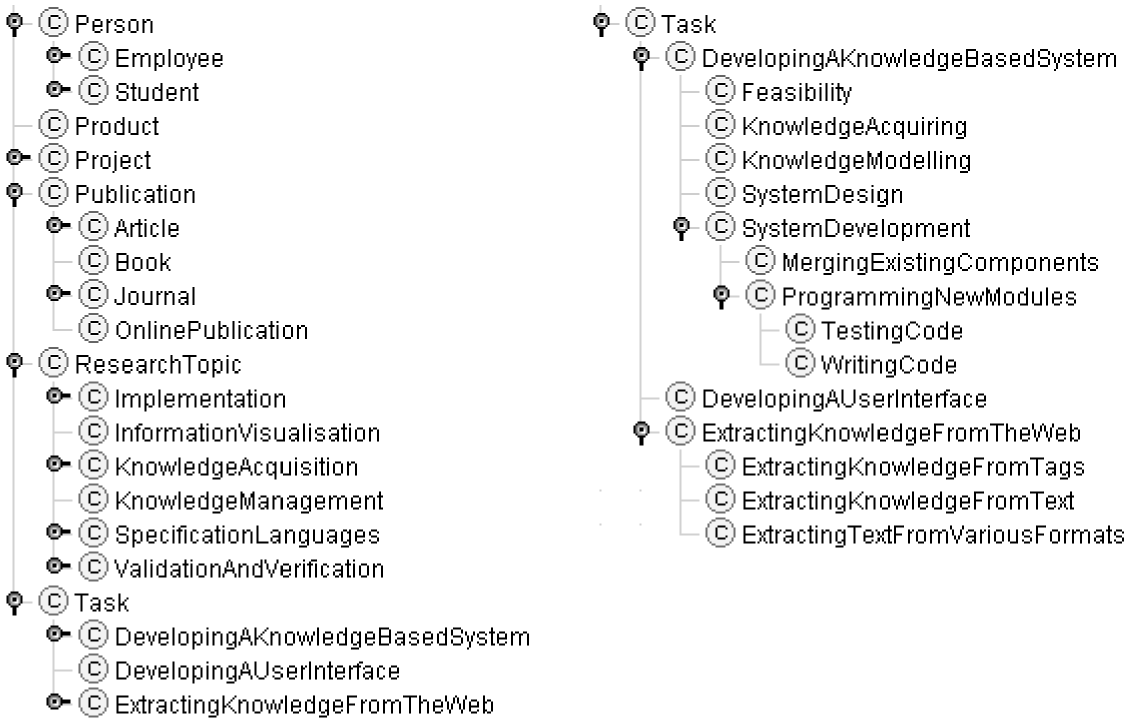

Fig. 2. Revised ontology showing "who", "where", "what" and "how" ontologies (left) and the current full ontology of tasks (right)

5 Ontology for Knowledge Valuation

However, we not only want to capture all the information about the knowledge present in a group of AI researchers, we also want to attach valuations to this knowledge. The reason is simple; even with a relatively small group of researchers, there may well be a problem of information overload because there is so much knowledge available about certain topics. The aim of knowledge valuation is to identify those items of knowledge that are (or are considered to be) more valuable.

Within the ontology, this is achieved by adding a couple of slots and an entire new ontology section. The slots are attached to the class of Publications; one is intended to record usage statistics (i.e. number of clicks for an online publication, or number of times borrowed for a paper publication), while the other records the "publication quality", which contains one or more terms from the new section of the ontology. These terms correspond to heuristics which give guidance on the quality of a publication; for example, "Few References", "Several References" or "Many References"; "One Reference Discussed

In Detail" or "More than One Reference Discussed In Detail"; and on the software side, "Prototype with One Example", "Proof of Concept system", or "Fully Functional system". An unlimited number of these entries can be attached to any publication (or project report), with the intention of being used as guidance on the quality of a publication.

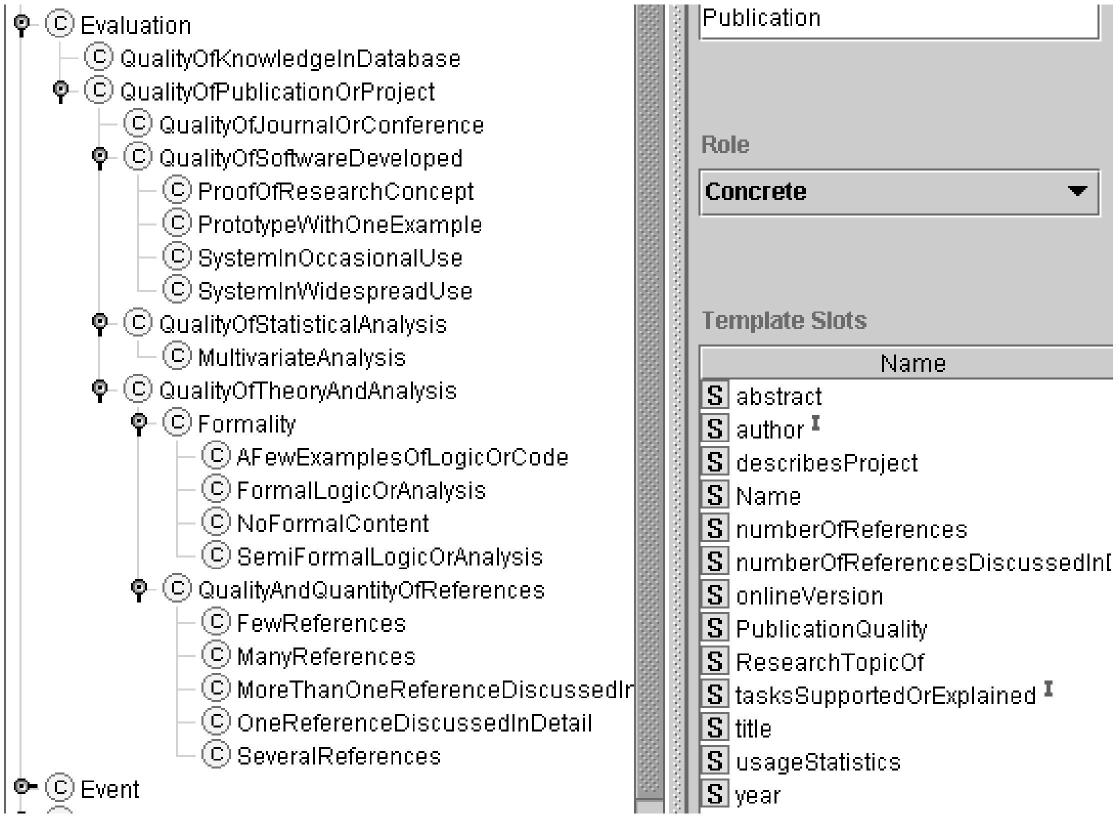

Fig. 3. Ontology for evaluation (left) and slots for Publication class (lower right). Note that several slots exist in order to collect valuation information

6 Future Activities

The priority in future work is to discover related research work and to use that to make further extensions to the ontology, either in the multi-perspective modelling area or in the valuation area.

We also have ongoing work on construction of a knowledge based system to help in determination of the quality of knowledge in databases. This system works by using heuristics to determine the likelihood of obtaining useful knowledge from the database with data mining techniques.

Another piece of work concerns the development of ontology meta-data to allow merging of knowledge models from different sources. This should be coordinated with the multi-perspective ontology work.

Acknowledgements

I would like to acknowledge the efforts of all my co-workers on the AKT project. I would especially like to acknowledge Stuart Aitken for guidance on existing valuation methods and for general advice.

This work is supported under the Advanced Knowledge Technologies (AKT) Interdisciplinary Research Collaboration (IRC), which is sponsored by the UK Engineering and Physical Sciences Research Council under grant number GR/N15764/01. The AKT IRC comprises the Universities of Aberdeen, Edinburgh, Sheffield, Southampton and the Open University.

The EPSRC and the Universities comprising the AKT IRC are authorised to reproduce and distribute reprints for their purposes notwithstanding any copyright annotation hereon.
The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing official policies or endorsements, either express or implied, of the EPSRC or any other member of the AKT IRC.

References

1. J. Kingston and A. Macintosh, "Knowledge Management through Multi-Perspective Modelling: Representing and Distributing Organizational Memory". Research and Development in Expert Systems XVI, proceedings of the Technical stream of ES '99, the 19th BCS SGES International Conference on Knowledge Based Systems and Applied Artificial Intelligence, 220-239, 1999.
2. J. Zachman, The Framework for Enterprise Architecture. http://www.zifa.com/zifajz02.htm.
3. A. Th. Schreiber, J. M. Akkermans, A. A. Anjewierden, R. de Hoog, N. R. Shadbolt, W. Van de Velde and B. J. Wielinga. "Knowledge Engineering and Management: The CommonKADS Methodology", MIT Press, ISBN 0262193000. 2000.
4. Knowledge Value-Added (KVA) Methodology Tutorial. http://www.iec.org/tutorials/kva/
5. Structural Knowledge Auditing http://www.nwaiag.com/what/struct.htm
6. KA2 - Knowledge Acquisition Community Ontology. http://ontobroker.semanticweb.org/ontos/ ka2.html
7. A. Preece, "COVERAGE: Verifying Multiple-Agent Knowledge-based Systems", Knowledge Based Systems, 12, 37-44, 1999.