2nd Hellenic Conf. on AI, SETN-2002, 11-12 April 2002, Thessaloniki, Greece,
Proceedings, Companion Volume, pp. 15-26

An Ontology-based Approach for Information Systems Development

Vaios Papaioannou[1] and Babis Theodoulidis [1]

[1] Department of Computation, UMIST, PO Box 88, Manchester M60 1QD, United Kingdom
tel. +44 161 200 3309, fax +44 161 200 3324
vp@mailbox.gr, babis@co.umist.ac.uk

Abstract. The wide use of *Ontologies* varies from fields such as knowledge management and natural language processing to enterprise modelling and information brokering. They provide a shared and common understanding of a domain that can be communicated between people and across application systems. The origins of ontologies can be found in Artificial Intelligence and some of the main aims are to facilitate knowledge sharing and reuse. The work reported in this paper is based on ontologies for supporting the early stages of *Information Systems development*. The approach presented, namely the *H.E.R.E. approach* targets the complexity that is associated with the process of *Requirements Engineering*. This framework is demonstrated with the development of both generic model and domain specific ontologies.

1 Introduction

The success of a system is measured by the extent to which it meets the needs and expectations of all stakeholders i.e. customers, users, developers. *Requirements* form one of the basic infrastructure elements for a system development process. They refer to the needs of the users and they include why the system is being designed, what the system is intended to accomplish, and what constraints are to be observed [18]. *Requirements Engineering (RE)* is regarded as one of the most important and critical phases of IS development because errors that occur during the early phases of an IS development, introduce much larger problems later on, during implementation, testing and maintenance [18], [3], [19], [2], [10], [12], [20].

The wide use of *Ontologies* varies from fields such as knowledge management and natural language processing to enterprise modelling and information brokering. They provide a shared and common understanding of a domain that can be communicated between people and across application systems. The origins of ontologies can be found in Artificial Intelligence and some of the main aims are to facilitate knowledge sharing and reuse [6] [9].

The work reported in this paper is based on ontologies for supporting the early stages of Information Systems development. The approach presented, namely the H.E.R.E. approach targets the complexity that is associated with the process of

Requirements Engineering. This framework is demonstrated with the development of both generic model and domain specific ontologies. Section two focuses on Requirements Engineering issues regarding our approach, the overview of our approach is presented in section three. Section four focuses on role of ontologies in the HERE project and section five presents some example ontologies.

2 H.E.R.E. and Requirements Engineering

In this section we present some of the issues of the RE process that are addressed with the HERE approach.

Requirements Engineering (*RE*) is defined as the systematic process of developing requirements through an iterative co-operative process of analysing the problem, documenting the resulting observations in a variety of representation formats and checking the accuracy of the understanding gained [13]. The aim of the Requirements Engineering process can also be viewed as getting from an initial input characterised as opaque personal views of the system represented using informal languages, to a desired output characterised as formally represented, complete system specification on which agreement was gained within the three dimensions of Specification, the Representation, and Agreement. [16].

During the RE process the RE stakeholders may be hindered by its complexity that may come into different forms such as *Product* complexity, *Organisational* Complexity and *Process* Complexity [15]:

More specifically within the scope of the HERE project we identified a set of issues concerning these complexities: *varying levels of information formality*, v*arying levels of information abstraction*, m*ultiplicity of representation formalisms*, m*ultiple views, multiple formats and mediums*, f*ragmentation of information.*

If RE is viewed as a problem solving process where all input and output information from the problem domain, determines the path from the initial input to the desired output, then the RE process is hindered from all the aforementioned complexities. On the other hand, within the HERE project this knowledge is encoded into a single, manageable, shared entity, an Ontology. The HERE approach is not radically new methodology for RE but a supporting framework that ensures a more secure path towards the desired output of the RE process.

3 The HERE Architecture

The information of a development project is multifaceted, voluminous, and comes in many different forms in order to cover all aspects concerning the problem domain. In HERE, we view this information essentially covering three main aspects of the Universe of Discourse:

- Knowledge about the *Domain requirements* and needs: enterprise requirements, user requirements, non-functional requirements, etc.

- Knowledge about the *Enterprise Domain*: the enterprise structure, its processes, actors, goals, etc.
- Knowledge about the *Information System*: system requirements, system design, database design, process design, etc.

This problem is addressed by the HERE approach both at the conceptual and the implementation level.

At the conceptual level, in the HERE approach we aim at two primary objectives:

- The integration and coupling of the disparate information elements which are produced during the RE processes into an explicit, unified, clarified, and defragmented representation. In HERE the above knowledge is explicitly expressed into sets of models and meta-models which describe concepts and meta-concepts respectively regarding the Universe of Discourse, in a modular, structured manner.
- The explicit specification of meta-models and models which describe the interconnections among these knowledge elements.
- In HERE, these two goals are accomplished with the use of *Ontologies*:

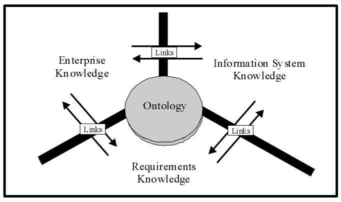

Fig. 1. : Integrating the problem domain knowledge in HERE.

At the implementation level, we aim at the management and communication of the above knowledge elements by providing tool support based on the HTML- World Wide Web standard. Figure 2 presents the overall architecture of the HERE environment.

In the HERE architecture, we identify three main components:

- The *Model Knowledge* component: Includes the knowledge that describes models that may be used in order to specify domains of interest. This information spans over he three levels of UoD knowledge as described earlier, i.e. Enterprise, Requirements, and System knowledge.
- The *Domain Knowledge* component: This component includes the knowledge that is modelled for a specific domain. The models are instances of model specifications in the Model Knowledge component. According to their content and purpose, they also span over the Enterprise, Requirements, and System knowledge.
- The *Repository* component: It holds all the above specifications in the form of *Ontologies* along with all relevant data that contribute to the overall knowledge of the UoD.

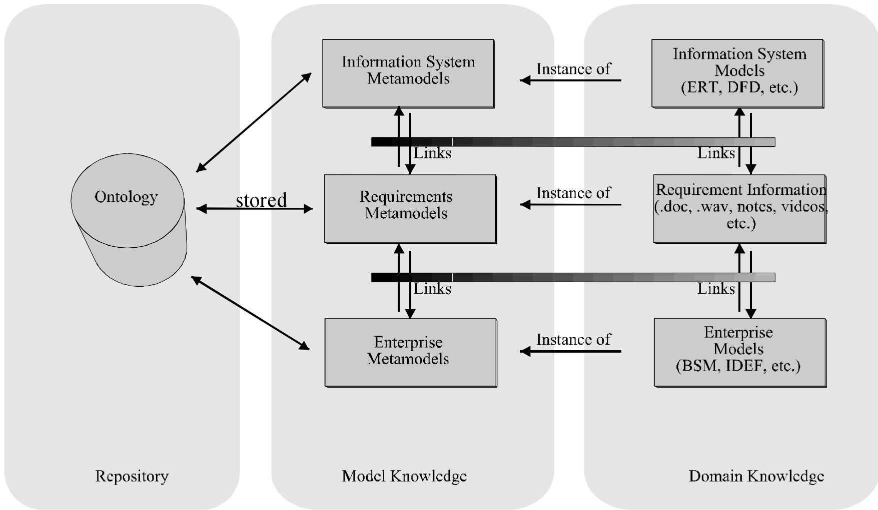

Fig. 2. : The HERE architecture.

4 The Role of Ontologies

One of the main priorities of the HERE approach was to address the problem of heterogeneous representations of all the information elements produced as a result of the RE processes.

The term *Ontology,* within our scope, is defined as: An explicit specification of conceptualisation. A conceptualisation is based on a body of formally represented knowledge; a set of definitions that allows one to construct expressions about some application domain [8], [24].

In our approach the concept of Ontologies was adopted within the context of *Requirements Engineering* and the main reasons are summarised below.

In HERE Ontology models differ slightly from traditional conceptual models. This is because Ontologies are driven from the need to describe the elements that exist in the UoD in a generic and explicit manner, so future references have to commit to these conceptualisations. Traditional conceptual models on the other hand are more specific in abstracting and describing the problem domain in the context of developing a system. In this sense, the content of Ontologies may be viewed as a superset of these models; for example, a specific data model of a problem domain may be based on the Ontology of the domain of interest. Thus, we view Ontologies as the ideal result of the RE process where there is an effective exploitation of the domain concepts and knowledge elements.

Another special characteristic of Ontologies is that they are explicit specifications of conceptualisations that are commonly agreed. Thus in HERE we view Ontologies within the RE context and regard them to include the agreement dimension:

Ontology = Model + Agreement

In this view Ontologies serve as more enhanced abstractions of reality which demonstrate a higher level of conformance to the three dimensions of requirements engineering, namely *specification*, *representation*, *agreement*, as described by K. Pohl ([16]) and may form a reliable basis of reference, communication, knowledge sharing and reuse.

Ontologies are inventories of the kinds of things that are presumed to exist in a given domain together with a formal description of the salient properties of those things and the salient relations that hold among them. In HERE Ontologies are used to hold the knowledge about models and model instances of the UoD in one unified representation. Thus by using a common representation ontologies become a computational medium to integrate all the domain information to describe domain knowledge in two primary realms: The *Metamodel* realm and the *Instance* realm:

The *Metamodel* realm; It is realised with the *model component* in HERE and it holds meta-level information, i.e. meta-models of the models that are used to describe the domain of interest. Examples are the ERT[1] model ontology, the Enterprise Ontology[2], etc. In this way, Ontologies form one single repository of domain knowledge that can serve as a reference point which can be reasoned throughout the development lifecycle. Apart from the model knowledge, in this realm we also describe the *Link Models*, which hold the information about the links that are pertinent to these models at the meta-level. Consequently, at the instantiation level, in the Domain component, the models automatically realise these meta-level links. Examples of Link models are the ERT to Enterprise intermodel links, etc.

The *Instance* realm; The instance realm is realised in the *domain component* of HERE where the meta-level ontology descriptions are actually instantiated with knowledge elements concerning the domain of interest. Examples are The ERT schema of the UMIST Admissions or the Enterprise Ontology of the UMIST Admissions department. The Model Link Ontologies are also instantiated at this level, providing the actual links that exist among the instance level information elements.

Another important point is that the Ontological descriptions are expressed in a *modular*, *expandable*, *reusable* manner. The fact that ontologies describe the model specifications used does not imply that the HERE ontologies replace these formalisms as the HERE framework provides support for the existing methodologies and formalisms and not the introduction of a new ones.

5 Ontology Development

Ontologies in HERE hold the knowledge concerning the problem domain and this includes all available information and artefacts produced during the RE process. This knowledge spans into two realms the meta-model realm and the model-instance realm.

---

[1] ERT stands for Entity Relationship and Time Model [21].

[2] The Enterprise Ontology is investigated by the Enterprise project[7].

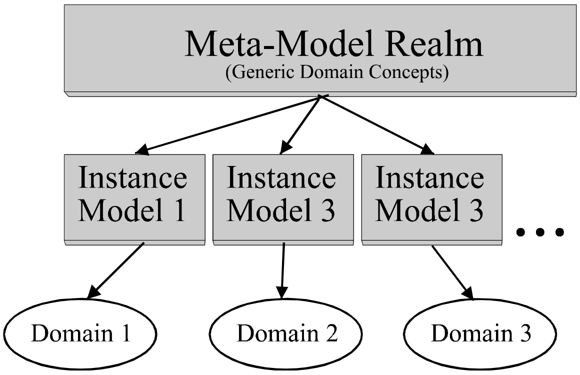

Fig. 3. : Levels of description in HERE.

The first step towards the development of an ontology for a specific model is the analysis of the main model concepts. Once the model is described as ontology at the meta-level then specific instances of the model can be used to describe individual problem domains.

In the HERE project ontologies are represented according to the object oriented paradigm. As it is understood that object oriented concepts are very common towards implementation, this must not be confused with our motivations. In our framework object concepts are mainly viewed as representation rather than implementation constructs. This involves objects into analysis and modelling of domain knowledge as opposed to implemented systems. In HERE the choice of representation relies on the object model as defined by the *ODMG* group [14] and more specifically the *Object Definition Language* (*ODL*).

In order to demonstrate our approach we have developed three representative ontologies each on covering one of the main areas of RE knowledge, Information System, Enterprise, and Requirements knowledge. We have also examined various link models and suggested a working set of link ontologies. Within the next section there will be a short presentation on how we specify ontologies.

5.1 The ERT Model

The *Entity Relationship Time*[3] model (*ERT*) is an extended entity-relationship model that additionally accommodates directly the representation of "time", as a distinguished entity, and "complex" objects. Besides the temporal dimension, ERT differs from the original E-R model in that it regards any association between objects in the unified form of a relationship, thus avoiding the unnecessary distinction between attributeships and relationships. The ERT model represents explicitly entity types and value types. For each relationship the ERT model recognises sentence

[3] This model is described in more detail in [21].

predicates which are used to make statements and referent functions which are used as a selection mechanism for entities or values.

After the analysis of the ERT model concepts then it is possible to code the facts concerning this model into an ontology. This ontology essentially consists of an object model hierarchy that can also be specified in ODL:

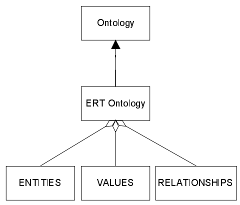

Fig. 4. : The ERT Ontology Hierarchy.

In ODL, we have the following:

```
interface ERT: Ontology
( extent Ontologies)
{
attribute set<ENTITY> Entities;
attribute set<VALUE> Values;
attribute set<RELATIONSHIP> Relationships;
};
```

The code above simply express the fact that the ERT model is consisting of Entities, Values, and Relationships.

Once the ERT definitions are stored at the meta-level this immediately allows the storage of information concerning specific problem domains as instances of these metamodels. For example let us consider a university admissions domain. Then we expect entity instances like students, admission officers, applications, etc. Then we an imagine an student entity object instance to be:

Object ENTITY

id: e1
full_name: student,
symbolic_name: student,
category: primitive
time_stamp: F,
concept_type: non_derived,
creation_date: 10/06/2001.
revision_date: 30/08/2001.
description: person who receives tuition,
author: vp

Similar instantiations may take place for the rest of the ERT concepts regarding the problem domain.

5.2 The Enterprise Ontology

The *Enterprise* project aims to provide a method and software support for enterprise modelling with a view to managing business change[7]. Enterprise models may be used by modellers, analysts, decision-makers, and operators. They may provide insight, aid communication or define enactment systems. The *enterprise ontology* modelling is intended to be used to guide modellers and application builders so that they use terms, which are consistent and commonly accepted. Then new models can be easily communicated to others who accept the ontologies and new models can be used for multiple purposes. New applications can be built which access, handle and present information in shared terms [23].

The original ontology was developed and coded within the scope and context of the Enterprise project. This includes the expression of the enterprise ontology in ontolingua [1]. Based on the HERE approach the Enterprise ontology had to be analysed and re-expressed in our representation language ODL. Our specification was coded based on the enterprise ontology informal specification ([23]) and supplemented by its ontolingua code counterpart ([1]). In our encoding of the enterprise ontology we introduced new assumptions, we included, modified or excluded original terms, and focused mostly on the enterprise ontology sections which were more relevant to our scope. The overview of our encoding of the enterprise ontology structure is presented in the following figure:

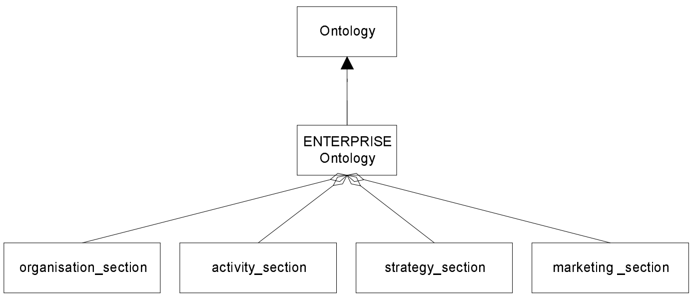

Fig. 5. : The Enterprise Ontology Hierarchy.

Essentially at the top level, we consider the enterprise ontology sections and then we refine these sections accordingly. Instantiation of the Enterprise Ontology may also take place in a specific domain like the admission domain as discussed above. For example we consider the ORGANISATION_UNIT instance Computation Department of the ORGANISATION_SECTION:

Object ORGANISATION_UNIT
 identity: computation_department
 managed_by: umist
 manages: (information_system_engineering, decision_ technology etc.)
 activities: set(admission, teaching, research, consultancy)

Similar to the development of these ontologies is also the *NFR framework* ontology [5] which is also included in the HERE set of encoded ontologies. More details about this set can be found in [15], [22]. The three ontologies together demonstrate our approach each one covering one specific aspect of the RE knowledge i.e. Information System, Enterprise and Requirements knowledge.

5.3 Link Ontologies

The conceptualisation of the relationships between the models and model constructs used to audit the problem domain knowledge, are as important as the elements of this knowledge themselves. In our case, the models and their links are expressed through ontologies. The link ontologies conceptualise the relations that exist either implicitly or explicitly between these ontology elements.

By expressing, all the knowledge of the universe of discourse (UoD) into one ontology consisting of several sub-ontologies is the one part of the HERE approach. Another task is to actually define the ontologies, which will link the various elements of this knowledge.

Link ontologies are modelled and defined just like any other ontology, as presented earlier. The essential difference between ontologies and link ontologies is that the latter emphasise on the relations among elements rather than pure element representation. Additionally at the implementation level, links are expected to be translated into hyperlinks according to our overall architecture.

Expressing links essentially involves the identity or type of the link, and ultimately its classification within its link ontology, a source element, and a destination element:

Link
Identifier
Source
Destination
Source to Destination Role
Destination to Source (inverse) role
Other additional properties

The links involve two nodes a source and a destination and two roles which reflect the two directions of the link, if the link is bi-directional. Additional properties may be added to the links such as attributes or even behaviour, expressed through object methods. Also it is possible for a link to target to more than one destinations if this is desirable.

In HERE the link ontologies are integrated with all remaining sub-ontologies contributing to the overall ontology. In HERE we may identify three categories of link ontologies, namely model link, traceability link, and user defined link ontologies (Figure 6).

By forming these link classifications and hierarchies, the HERE users are able to fully represent their knowledge concerning links among knowledge elements into clear, well structured, comprehensive, explicit models. The same models can be reused into similar situations, and tailored according to project needs. New link models or inappropriate models may be included and excluded on demand. In addition, certain stakeholders may focus on certain link models; e.g. software developers may select and reason about the links concerning the information system – software models.

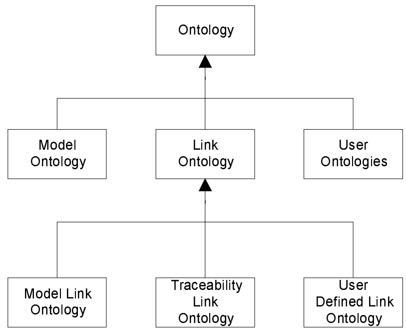

Fig. 6. : The Overall HERE Ontology Hierarchy.

The refinement of the top-level ontology diagram presented above depends on the very domain that all these models will apply. In our analysis some main classifications of links were studied. For example, in refining the Model links we may consider the existence of intermodel an intramodel links:

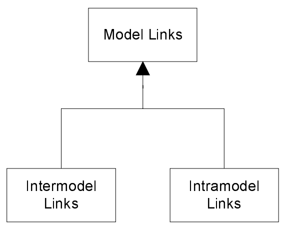

Fig. 7. : The classification of Model Links in HERE.

So if we consider a model concerning intermodel links such as the F^3 model ([4]) that expressed the links between the various "worlds" in Enterprise Modelling we may come up with the following object diagram:

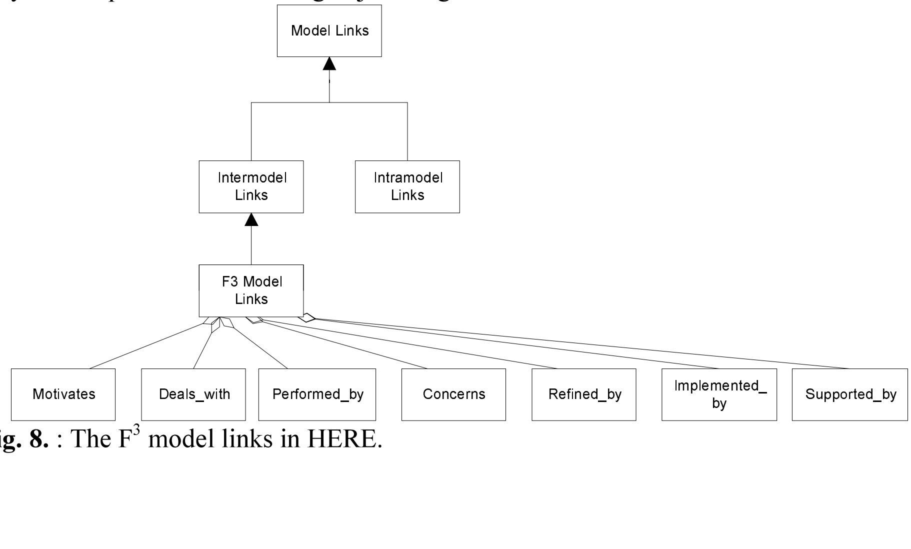

Fig. 8. : The F^3 model links in HERE.

Similarly we have considered other link models such as the *HYDRA* taxonomy of *referential link types* ([17]).

6 Conclusions

The HERE project is critically based on ontologies. It addresses the complexity of the RE process by supporting existing practices through a complementary framework. This is achieved with the effective management and co-ordination of the information gathered throughout the early stages of a system development into a single consistent manageable entity.

Our ontology framework allows sharing and communication of information since all the parties involved can share this common point of reference. Moreover, the ontology models can be reused either partially or as a whole to future situations. We advocate the use of models which underline the notion of agreement. Agreement is one of the main dimensions in requirements engineering and this comes as an addition to the already supported representation and specification dimensions.

HERE attempts to provide logical links amongst models, designs, requirement specifications, and other informal information and thus, it allows reasoning for the impact of one part of the system development knowledge to another and the correlation between specifications (possibly formal) and their actual (possibly informal) sources.

Current and future work for the HERE project among others include the development and integration of new ontology models and link models, the possibility of improving the ontology representation e.g. adopting the T.A.U. model [11] in order to benefit from time reasoning and schema versioning or try to incorporate some intelligence e.g. case-based reasoning for the stored models.

References

[1] "Ontology ENTERPRISE-V1.0, Ontolingua code" Artificial Intelligence Applications Institute, 80 South Bridge, Edinburgh EH1 1HN, Scotland, UK, 8 November 1996.

[2] ARIANE 5 Flight 501 Failure Report by the Inquiry Board The Chairman of the Board : Prof. J. L. LIONS E.S.A. (European Space Agency) Public Relations Press Releases and Information Notes, Paris, 19 July 1996.

[3] Boehm, B.W. "Software Risk Management"' IEEE Computer Society Press: Washington, 1989.

[4] Bubenko Janis A. jr., "Extending the Scope of Information Modelling" Fourth International Workshop on the Deductive Approach to Information Systems and Databases, Lloret, Costa Brava (Catalonia), Sept. 20-22. 1993.

[5] Chung Lawrence "Representation and Utilisation of Non-Functional Requirements for Information System Design"

[6] D. Fensel. Ontologies: Silver Bullet for Knowledge Management and Electronic Commerce. Springer-Verlag, Berlin, 2000.

[7] Fraser, J. (1994) "Managing Change through Enterprise Models" In Milne, R., Montgomery, A. (Eds) Applications an Innovations in Expert systems II, SGES Publications, 1994.

[8] Gruber T. "A translation approach to portable ontology specifications" Knowledge Acquisition, 5,pp199-220, 1993.

[9] I. Horrocks, D. Fensel, C. Goble, F. Van Harmelen, J. Broekstra, M. Klein, and S. Staab. "The ontology inference layer oil". Technical report, Free University of Amsterdam, 2000. http://www.ontoknowledge.org/oil/.

[10] Succeedings of the 8th International Workshop on Software Specification and Design (IWSSD-8), Schloss Velen, Germany, 22-23 March 1996.

[11] Kakoudakis, I. and Theodoulidis, B. "TAU: Towards a Unified Temporal Information Management Framework", Bulletin of the Italian Association for Artificial Intelligence, (AI*IA), March 2001.

[12] Report on the London Ambulance Service (LAS) inquiry, February 1993, version 0.99, electronically prepared by Antony Finkelstein, 30-11-95, original ISBN no: 0 905133 70 6.

[13] Loucopoulos P., Champion R.EM. "Knowledge Based Support for Requirements Engineering" Information and Software Technology 31(3), pp. 123-135, April 1989.

[14] Edited by R.G.G. Cattell, Douglas Barry, Dirk Bartels, Mark Berler, Jeff Eastman, Sophie Gamerman, David Jordan, Adam Springer, Henry Strickland, and Drew Wade "The Object Database Standard: ODMG 2.0" Morgan Kaufmann Publishing, ISBN 1-55860-463-4, 1997.

[15] Papaioannou V., "HERE: Hypermedia Environment for Requirements Engineering" Ph.D. Thesis, Department of Computation, University of Manchester Institute of Science and Technology, Manchester, U.K, 1998.

[16] Pohl K., "The three dimensions of requirements engineering: A Framework and its applications " Information systems Vol. 19, No. 3, pp. 243-258, 1994.

[17] Pohl K., Haumer P. "HYDRA: A Hypertext Model for Structuring Informal Requirements Specifications" Proceedings of the Second International Workshop on Requirements Engineering: Foundations of Software Quality, REFSQ '95, Jyvaskyla, Finland, 12-13 June, 1995.

[18] Ramamoorthly C.V., Prakash A., Tsai W.T., Usuda Y. "Software Engineering: Problems and Perspectives", IEEE Computer, pp. 191- 209, October, 1984.

[19] "Systems Engineering Capability Maturity Model, Version 1.1, A" Software Engineering Institute report CMU/SEI-95-MM-003, November 1995.

[20] "CHAOS" The Standish Group International, Inc., http://www.standishgroup.com/chaos.html, 1995.

[21] Theodoulidis, C., Loucopoulos P., Wangler, B. "A Conceptual Modelling Formalism For Temporal Database Applications" Information Systems, 16(4): p401-416, 1991.

[22] Papaioannou V. "The HERE Ontologies" TimeLab Technical Report TR-98-1, Department Of Computation, UMIST, 1998.

[23] Uschold M. "Towards a Methodology for Building Ontologies" Workshop on Basic Ontological Issues in Knowledge Sharing, held in conduction with IJCAI-95, 1995.

[24] B. Wielinga, A. Th. Schreiber, W. Jansweijer, A. Anjewierden, and F. van Harmelen. "Framework and formalism for expressing ontologies." ESPRIT Project 8145 KACTUS, deliverable DO1b.1, UvA, 1994.