

NOPIK: Creating Shared Conceptualisations for Information and Knowledge Sharing in the Virtual Enterprise

Albert Bokma, Norman Solomon, Howarth Harvey & Susan Jones

*School of Computing and Technology, University of Sunderland, St. Peter's Campus,
DGIC – Room DM401, Sunderland, SR6 0DD, United Kingdom, albert.bokma@sunderland.ac.uk*

Abstract

IT supported ways of working are indispensable when working in a virtual enterprise where there are geographic separations to be overcome and where the effective and managed sharing of resources is indispensable to its efficient operation. Current web technologies have significant limitations with respect to communicating not only resources but also semantics to enable a much more context sensitive processing, presentation and navigation of resources. This paper reports about progress in the development and evaluation of an ontology based tool that allows individuals and groups to easily develop shared conceptualisations and to use these conceptualisations to index shared resources. The paper is presenting a user managed application of lightweight ontologies to information and knowledge sharing which is currently being evaluated with a number of enterprises. The work has been done as part of the NOPIK project (IST-2001-33487) where the tool presented here is a part of a more comprehensive set of tools.

Keywords

Ontologies, ontology interfaces, information management, information navigation

1 Introduction

When people work together their shared understanding and conventions help them share and correctly interpret artefacts such as data, files and other records. It is difficult to reproduce this in a more flexible work environment and what can otherwise be relied upon as tacit and common understanding will have to be made explicit to enable effective sharing. A number of tools such as portal solutions have emerged to support the better sharing of information resources in an organization. This includes advanced indexing solutions such as Portal-in-a-Box™ by Autonomy, to name but one, but which on the whole do not allow users to create their own shared conceptualizations nor store and browse content flexibly, based on a domain model, rather than linguistic principles. While general purpose solutions such as these doubtlessly have their uses, we argue that more domain specific solutions will better support the more focused activities either in a single organization and more so still in virtual organizations.

What is needed are much more semantic based approaches that will facilitate the work of the individual, enabling them to cope with large information collections effectively and to be able to selectively share resources with others in a collaborative setting. IT supported ways of working are indispensable when working in a virtual enterprise where there are geographic separations to be overcome and where the effective and managed sharing of resources is indispensable to its efficient operation. While the current web technologies provide considerable scope for solutions that allow easy access to shared information, there are nevertheless significant limitations with respect to communicating not only resources but also semantics to enable a much more context sensitive processing, presentation and navigation of resources.

The semantic web aims to remedy these shortcomings with the use of ontology based sharing of conceptualisations and indexing of resources. There is considerable scope in the application of

ontologies and the semantic web to the problem, but which needs to be harnessed to deliver results for the individual user and user communities in particular, such as in virtual enterprises.

In this paper we briefly highlight certain functions of the application. This is followed by an evaluation of the application with respect to information systems and their role in the virtual enterprise.

2 The Cognet Application

Cognet is an ontology based information management application developed as part of the NOPIK project (IST-2001-33487). The purpose of this application is to enable users to create and use lightweight ontologies for the purpose of storage, navigation, retrieval and sharing of documents. The application uses primarily a graphical interface (see figure 1) where the ontologies are represented as a *concept maps*. The maps represent a domain specific taxonomy and which are visualised as a (directed acyclic) graph and can be inspected and navigated by the user. The screenshot below shows the main graphical interface. The user can explore the concept maps and documents that are attached to them and interactively edit them. A user management system regulates user access to concept maps and attached documents in a multi-user environment.

The system is implemented as an Internet web application using PHP and JAVA technology, which stores ontologies created by users in a MySQL database which resides on a web server. A Java applet is downloaded to the web client and provides the interactive graphical interface. This allows users to build up libraries of ontologies which can be selectively shared across the Internet. Ontologies can also be imported and exported in RDF(s) format (with capabilities for supporting DAML+OIL and OWL formats currently under development). The graphical interface/editor has been designed primarily for ease of use, and its implementation has been guided by usability testing.

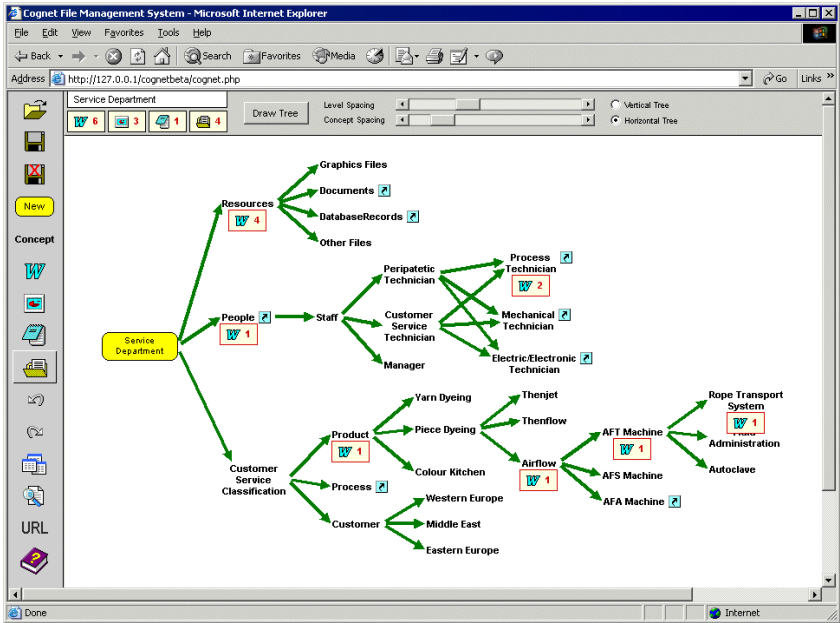





Figure 1: The Cognet Interface

Ontologies are displayed as concept hierarchies. The arrows connecting concepts are inverse to those used conventionally in ontology visualisations as this was found to be more intuitive to the non-expert user. We have also found that users tend to prefer graphical representations such as trees, rather than tables, when visualizing domains in their mind's eye; tabular concept

hierarchies as used in tools like Protégé [Protégé, 2004] do not easily allow users to appreciate the totality of a taxonomy and how concepts fit together.

Ontology paging in Cognet allows users to break up large ontologies into separate “pages”. In the example the four pages shown on the left have been merged automatically by Cognet to form the ontology on the right. The user can navigate the paged structure by clicking on the small arrows next to the “link concepts” **BBB**  **CCC**  **XXX** . These concepts are used as ontology navigation devices, and as such they are not part of the logical concept hierarchy and are removed during the merging process. Cognet also allows users to automatically merge the pages that make up a single ontology, as shown above. This merging algorithm can also be used to merge ontologies of arbitrary size, and it detects and merges (combines) duplicate concepts automatically. For example, the user could import a new ontology and merge it with an existing ontology. If the merged ontologies both contained the same concept, then these concepts would be combined to form a single concept within the merged ontology.

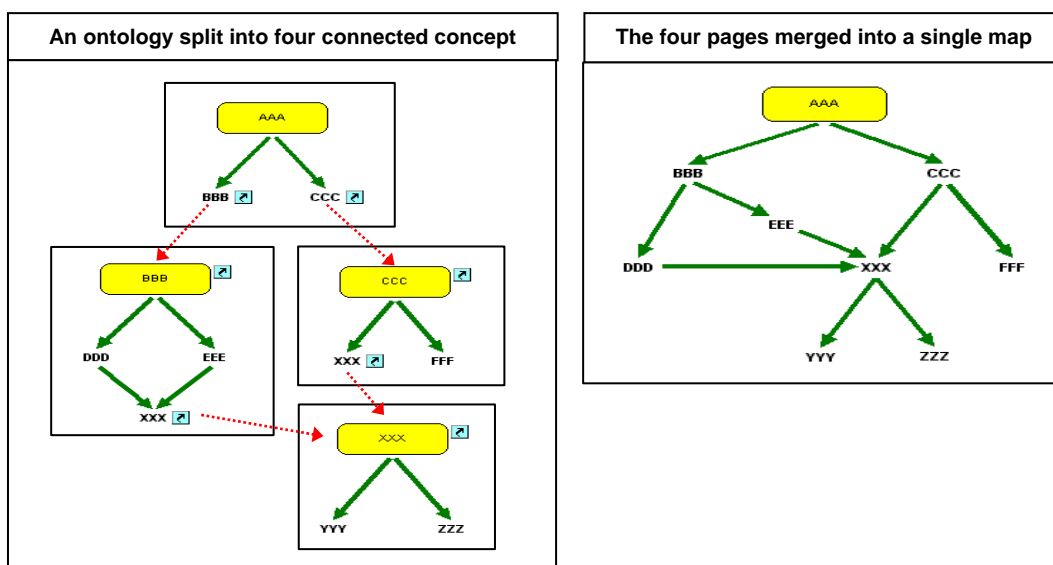


Figure 2: Ontology Paging in Cognet

Usability testing has also demonstrated that users prefer not to scroll large ontologies on screen. The Cognet system solves this problem by providing a paging mechanism which allows users to break up large ontologies into “screen size chunks”, as shown in Figure 2 but which, though represented to the user as separate maps are in fact logically connected in a single structure which is generated for the purpose of exporting the ontology in RDF(s) format.

2.1 Concept Map Editing Functionality

Cognet provides several automated ontology tree drawing options which allow users to automatically arrange and redraw the concept hierarchy in a tidy layout. Cognet also provides several manual ontology editing functions which allow users to edit and perfect the structure and the layout of the concept hierarchies. This functionality augments the automatic layout options described above, and is consistent with the semi-automatic nature of the proposed tool.

Child concepts can be added and automatically connected to any selected parent within the class hierarchy with a single click of the mouse. Also concept to concept arcs (arrows) can be added, repositioned or removed using a click and drag mechanism. This allows users to edit the class hierarchy by adjusting the arcs connecting concepts. Also the sub-tree below any concept can be automatically converted (clipped) into a separate concept map with a single mouse click. This allows users to easily create and edit the concept map structure.

2.2 Document Classification and Browsing Functionality

Cognet allows the user to classify and browse documents through the graphical interface. Thus a document can be associated with a number of concepts that represent the content of the document or its purpose, which overcomes the typical dilemma of classical filing hierarchies where a document can only be stored in a single place under a single organising principle; this is usually represented by the logic of the filing hierarchy.

An automatic classifier helps users attach new documents to a given set of concept maps based on matches between concepts on the maps and their occurrence in the documents. The user can also change these associations manually.

In order to locate documents on the maps the user can either click on the relevant button on the toolbar to see the location of documents of that type as shown by small pop ups appearing below the concepts with a number indicating the number of documents found for that concept. Alternatively, the user can hover with the pointer over a concept to see the number of associations of any type as shown by a sample set of concepts on the left in figure 3 and where the associations for one of the documents generated by the classifier are shown in the relevant pop up window on the right.

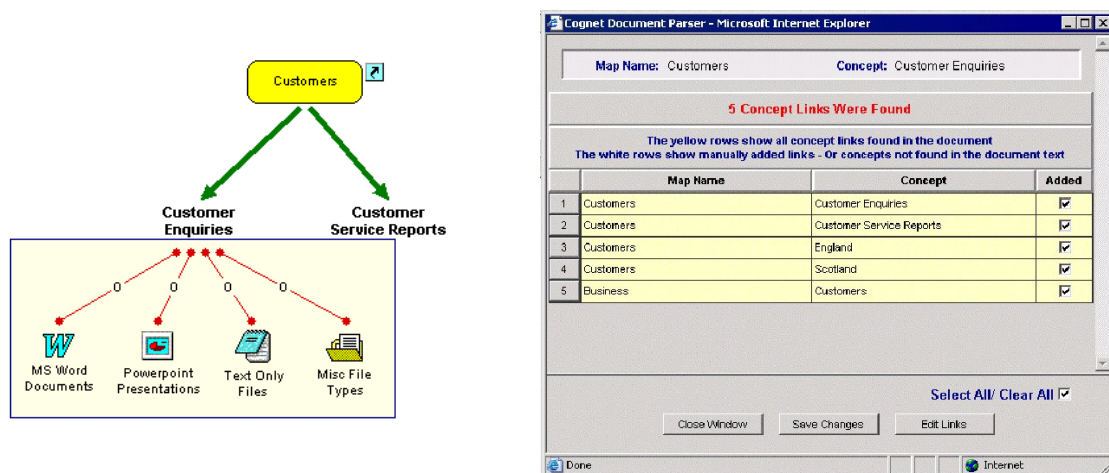


Figure 3: Document Association and Classification

3 Evaluation Results

In the context of the NOPIK project the evaluation of the Cognet application was carried out in two distinct strands:

- Partners of the project consortium, *project partner users*, that are users of the technology have investigated the application in a quasi case study/action research style
- An additional group of *voluntary users* has been recruited to look at the application at a conceptual level to assess its potential in less in-depth fashion, but within a wider range of contexts than the project partner user group

The larger group of *voluntary users* consisted of 19 people from 10 organisations that were selected from the voluntary, commercial and public service sectors, where the total contact time was approximately a day and a half. Following an introduction to the concepts of the application and a structured interview on their use of IT, they were invited to see the application and engage in an exercise in its deployment through an example. The example involved going through a categorisation process around an activity which was considered to be common to all. The volunteers were only given some rudimentary rules of on how to construct an ontology. Having manually constructed an ontology, they transferred this onto the Cognet application, uploaded documents and saw how they became attached and modified the links as they saw fit.

By contrast, in the case of the *project partner users*, a concerted attempt was made to develop a methodology for the implementation of Cognet by the users with the objective of minimising the use of outside expertise. Initially, this methodology involved the application developers carrying out a case study/action research based use case. Subsequently, a more hands-off approach was adopted in prototype deployment with the *project partner-users* in order to understand how the methodology might work.

3.1 Methodology for Implementing Cognet

The methodology recognises that distinct working groups in an organisation may have the different perspectives about the same process (e.g. doctors and nurses in relation to patient treatment) or how different activities around a given artefact may produce different perspectives on that artefact (e.g. designers and maintenance staff in relation to a particular machine). It therefore emphasises the need to identify the scope of the network of people who can work collaboratively to establish common meaning to the ontological concepts and their map structures within Cognet.

The methodology does outline some strategies for identifying concepts and their labels. Since the application is about storage and re-use of information within documents, there is a need to look at the information assembled in existing documents. The rules for building ontologies would generally still be considered very rudimentary by those who are creating ontologies for more general purpose applications like machine translation [Uschold and King, 1995]; [Welty and Guarino, 2001 and 2002], particularly with regard to more sophisticated relationships between concepts. Our emphasis has been on local ownership and local 'common sense', a tool for self-building and self-evolution of ontological structures. We also suggest there needs to be relative stability and a clear organisational mechanism for changes to the ontologies. At the start it is anticipated that there will be a number of iterations before a stable structure is achieved.

3.2 Points Raised by Users

There was a general recognition amongst the larger *volunteer user group* that the application solution is addressing a real problem within their organisations. This group clearly felt at ease with transferring their manually constructed ontologies into the application. The ease with which it is possible to draw a structure and manipulate it, upload and download documents was recognised by all the participants, even those with quite limited experience of general purpose office applications.

Although there were no virtual organisations participating, those who belonged to distributed organisations immediately saw the benefits of having a networked application structure that permits joint working on common documents and other elements of sharing. Even in such a short period of familiarisation, it was recognised that the categorisation or conceptualisation of information and knowledge was complex.

The *project partner users* had more time to develop maps and have gone through more iterations in developing their ontologies. They have managed to avoid some of elementary errors like inconsistent degree of decomposition by the larger volunteer user group. Repetitive uses of concepts labels have normally been coped with by creating independent trees or re-labelling. There have been requests for some quantitative scaling between levels of decomposition and temporal measurement between instances and/or grading of instances.

3.3 Discussion

Two important aspects about local information/knowledge need to be raised:

- how the application constrains its representation plus issues of temporality of local information or its perceived 'weight'.

- how autonomous this local information should remain with respect to the overall organisation

Looking at the first issue, it is interesting to observe that although they are coming from a slightly different direction, [Adams & Avison, 2003] in talking about impact of technique on the representation of problems, attempt to characterise techniques in relation to a matrix which has coordinates of closed to open and prescriptive to non-prescriptive. Hierarchical techniques tend to come at the closed, prescriptive end.

With respect to Cognet the end user is able to choose any number of hierarchies and decompose them to a self-chosen step size and depth. Nevertheless, they remain hierarchies and one constraint is the more or less singular relationship indicated between any parent and child concept. This is clearly distinct from a human parent-child relationship which can be multifaceted, each facet or combination of facets being weighed differently at different times and contexts [Pentland, 1995] and [Geertz, 1983].

In some cases, it is possible for individuals to have multiple points of view with respect to nominally the same information. If there is a need to act on this information, then a choice will be made on a singular perspective dependent on the time and context. Clearly, there are other examples in work settings where a singular relationship that is unambiguous can be seen as desirable. Additionally, there are times when compromises in richness of information are useful.

Considering the question of whether the application places constraints on the representation of local knowledge, raises the issue of how people might represent it without any application. In many work situations, it would seem reasonable to say that staff are not asked to do this. It is a somewhat reflective task and is likely to require some kind of facilitation. A number of issues arise from this not least of which are certain temporal aspects. At any given point in time, people are likely to be heavily influenced in their perspective on the recent past. This also arises in explanation of deficiencies in the analysis of organisational decision making [Schwenk, 1985], [Einhorn, 1982] and requirements definition [Haumer et al., 1999].

To some extent this can be overcome by putting the application through a number of iterations over a period of time before any ontology is considered to be stable. A second temporal issue is the degree to which the framework of local knowledge changes over time. If this is significant, then also the potential for re-use of local information may also be limited. Alternatively, if there is no structure, will the possibility of interpreting change not also be constrained? Within Cognet at least there is the possibility of local (end user managed) control over the evolution of the ontologies. As the tool supports multiple ontologies that can exist and be used in parallel there is also the possibility of associating the same document with separate ontologies that represent different perspectives.

With respect to the second main point of how autonomous this local information should remain, it has to be added that there are times when information/knowledge has to pass out from the local [Vincenti, 1990] and [Egyedi, 2001]. However, for much of the time this is not the case and although collaboration is needed between groups of different perspectives within an organisation, it can be quite destructive to subsume their perspectives under a single regime. It can destroy meaning for both parties and condemn the structure to limited use – this is a known issue for example in information systems that have to support a variety of uses and users [Monteiro and Hepsø, 2002].

A very specific feature of Cognet is to allow local control of the conceptual structure of information used in a given locality and on the whole this has been an under-recognised alternative in information system development. Arguments for ‘localised universals’ have been presented by [Hartwood et al., 2001], in the development of an ‘Electronic Patient Record’ for use in hospitals. They argue that this allows for decisions about information needs of potential remote users to be deferred to the latest possible time. Translations of local information will still be required but they involve specific ‘contracts’. People at the source of information can remain

aware of who uses their information and in what ways, i.e. an intelligent translation can be negotiated at the time of exchange where this is required. [Hartswood et al., 2001] (op.cit.) argue for a combination of 'global universal' information, e.g. patient identity and 'local universal' information, information specific to a particular section or profession within the hospital.

A different kind of localised information arose within a previous project use case investigated by the authors [Slade and Bokma, 2002] involving a manufacturer of textile dyeing machines. Here the more global information of the specification of the dye chemical had to be transformed for deployment in their machine types (to some degree, chemistry dynamics are specific to a machine type) and from time to time to the specific context of the user (specific chemical/textile combinations and machine installation). Responsibility for the 'recipe' for dyeing a textile or resolving dyeing defects is often the subject of individual, though not necessarily unique solutions. The tendency has been for the machine designer to adopt a philosophy of acquiring repeatability through ever more sophisticated systems of control rather than universalising the understanding of the complexity of the chemical dynamics in a given machine type.

Clearly, this does not exclude the possibility of gaining a more universal understanding of the dyeing process for their equipment and this will in part be facilitated by the dyeing technicians sharing their knowledge. It is anticipated that the Cognet application could contribute to this. However, at this time it was considered inappropriate to try and incorporate the knowledge of the dyeing technician, machine designer and dye manufacturer into some greater universal structure.

In general, the local and the global are working together. Our argument is that the local universal has been under-represented in information systems and that there is an argument for facilitating transmission and negotiation of information between local universals. Cognet does in fact support both uses as a single ontology shared by all could be used to represent the universal perspective, while the local perspective can be maintained by the parallel use of local or even personal ontologies for the same collection of documents or artefacts.

4 Implications for the Virtual Enterprise

The Cognet tool allows individuals and groups to easily develop shared conceptualisations in a graphical way and to use these conceptualisations to index resources that can be adjusted by the user. The work reported here is pioneering the user managed application of ontologies to information and knowledge sharing and comprises a tool and a method for its application.

A networked application designed for joint sharing and working on online documents, like Cognet, clearly serves a purpose in the virtual enterprise, as it would in any distributed enterprise or an enterprise with mobile staff or staff that are teleworking. An important issue in this scenario is the notion of sharing and whether there is a difference or similarity in the meaning of sharing within the virtual enterprise relative to other types of organisation. A virtual enterprise exhibits boundaries across which information has to be passed and is likely to represent a boundary of difference in expertise as well as in organisational structure. Moreover, there can be a significant difference between those virtual enterprises that have been brought together to develop a new product requiring a synergy in innovative ideas and those virtual enterprises which are put together to produce an artefact which has been largely specified. The alternative would be to perceive the virtual organisation as a collection of local universals where information is subject to certain transformations through negotiation at or near the point delivery to another partner in the organisation.

There is an argument for adding to Cognet the capability to manipulate document collections or part documents in order to meet the requirements of at least some types of virtual organisations. Thus instead of exporting ontologies, functionality could be introduced to manipulate information residing in one conceptual framework in the process of its transmission to a second or additional framework residing in partner organisations.

The Cognet tool already supports the parallel use of ontologies vis-à-vis a common set of documents. The classification mechanism supports the classification of a new document against all the ontologies available to a particular user and where each user can have as many ontologies as is desirable. In this respect the tool can support applications where different users use different ontologies (especially where they cannot or will not agree on a common ontology). At present the tool is hosted on a web server where multiple users use a shared installation. In the final phase of the project enhanced sharing facilities and the support of sharing between installations is envisaged. In the course of the evaluation it became apparent that there are situations where the users were not able to use an application hosted elsewhere, which was to host sensitive documents, thus prompting the developers to look more carefully at sharing management issues and communication issues between distributed installations, as would also be typical of a virtual enterprise which comes together only for a limited period and where privacy and non-proliferation of key commercial information and documents are an issue. While the Cognet tool does not solve the totality of the problems faced, it nevertheless has made an important contribution to evaluating the real issues in applying ontologies as an organising principle under the control of the end user and with favourable results so far.

Acknowledgement

This work has been partly funded by the European Commission IST Programme. The authors wish to gratefully acknowledged their support as well as the contribution of the NOPIK project partners to the evaluation of the tool.

References

- Adams, C and Avison, D: "Dangers inherent in the use of techniques: identifying framing influences", *Information Technology and People*, Vol.16, No.2, 2003, pp 203 – 234.
- Egyedi, T.: "Infrastructure flexibility created by standardised gateways: The cases of XML and the ISO container", *Knowledge, Technology and Policy*, Fall, 2001
- Einhorn, H.J. : "Learning from experience and suboptimal rules in decision making", Chapter 19, pp 268 – 283 in "Judgement under uncertainty: heuristics and biases", editors Kahneman, et al, Cambridge University Press, 1982
- Geertz, C: "Local knowledge: further essays in interpretive anthropology", publisher: Basic Books, 1983
- Hartwood, M; Voß, A; Procter, R; Rouncefield, M; Slack, R and Williams, R: "Building Information systems as universalised locals", *Knowledge, Technology and Policy*, Vol.14, No.3, pp 90 – 108
- Haumer, P et al: "Bridging the gap between past and future in Requirements Engineering(RE): a scenario-based approach", in *Proceedings of IEEE International Symposium on Requirements Engineering*, 7 – 11 June, 1999, Limerick, Ireland.
- Monterio, E & Hepsø, V.: "Purity and danger of an information infrastructure", *Systemic Practice and Action Research*, Vol.15, No.2, April, 2002, pp 145 – 167.
- Pentland, B.T: "Information systems and organizational learning: the social epistemology of organizational knowledge systems", *Accounting, Management and Information Technology*, Vol.5, No.1, 1995, pp 1 – 21.
- Protégé-2000 Home Page (Information and free download), Available at: <http://protege.stanford.edu>
- Schwenk, C.R.: "The use of participant recollection in the modelling of organisational decision processes", *Academy of Management Review*, Vol.10, No.3, 1985, pp 496 – 503
- Slade A. & Bokma A.: "Ontologies and XML", *Proc. of the 35th Annual Hawaii Int. Conf. on System Science*, IEEE Computer Society Press, Jan 2002, ISBN: 0769514359
- Uschold, M & King, M: "Towards a methodology of building ontologies", *AIAI-TR-183*, July, 1995
- Vincenti, W.G.: "What engineers know and how they know it", The John Hopkins University Press, ISBN 0-8018-3974-2
- Welty, C & Guarino: "Evaluating ontological decisions with Ontoclean", *Communications of the ACM*, Vol.45, 2002, No.2, pp 61 – 65.
- Welty, C & Guarino: "Supporting ontological analysis of taxonomic relationships", *Data and Knowledge Engineering*, Vol.39, 2001, pp 51 – 74.