Assessment of Ontology-based Tools: A Step Towards Systemizing the Scenario Approach

Alain Giboin, Fabien Gandon, Olivier Corby, and Rose Dieng

INRIA Sophia Antipolis, Acacia Project,
2004 route des Lucioles, B.P. 93,
06902 Sophia Antipolis Cedex, France
{giboin|gandon|corby|dieng}@sophia.inria.fr
http://www.inria.fr/recherche/equipes/acacia.en.html

Abstract. Scenarios have been already used for designing and evaluating ontology-based tools. For example, the so-called "motivating scenarios" are a core component of the TOVE ontological engineering method elaborated by Grüninger and Fox (1995; Uschold and Grüninger, 1996). We ourselves used a "scenario approach" for designing and evaluating CoMMA, a corporate memory computer platform based on ontologies and agents; the approach was inspired by the scenario approaches proposed in the HCI and CSCW communities, which we consider more user-oriented than the "motivating scenarios" approach. In this paper, we account for our CoMMA experience and its major lesson: the *necessity to apply the scenario approach more systematically* for assessing the usability and utility of ontology-based tools.

1 Introduction

In *"Some Ideas and Examples to Evaluate Ontologies,"* Asunción Gómez-Pérez [17] made the distinction between evaluation and assessment of knowledge sharing technology (KST), which include ontology-based tools (OBTs): "Evaluation means to judge technically the features of KST, and assessment refers to the usability and utility of KST in companies" – more precisely, as stated elsewhere by Gómez-Pérez [18], assessment refers to "the usability and utility of the ontologies, software environment, and their documentation when they are used within a given organization or by software agents."

The ontology community seems to be more concerned with the evaluation of OBTs, and with providing technological evaluation criteria such as interoperability, "turn around ability," performance, memory allocation, scalability, or integration into frameworks (see, e.g., [1]). In this paper we will rather focus on assessment of ontology-based tools, and on providing usability and utility criteria motivated by scenarios of use. Why? Our own experience of OBT design makes us think that we would not neglect assessment if we want to get "a consistent level of quality and thus acceptance" of OBTs by industry.

A way to give its place to assessment, we claim, is to make a more systematic use of user-centered scenarios, or to apply a scenario approach more systematically. This claim rests on, and is a major lesson of, our experience of the design and assessment of CoMMA, a corporate memory computer platform based on ontologies and agents [6][12][13], and on Corese, a semantic search engine designed in our research team [7][8]. In this paper, we will account for our CoMMA experience, and introduce some considerations about the systemizing of the scenario approach to ontology-based tool design and assessment.

2 Limitation of the "Motivating Scenarios" Approach Familiar to the Ontology Engineering Community

From February 2000 to January 2002, we participated to the IST European CoMMA project aimed at designing the CoMMA platform. The CoMMA project gave us the opportunity to apply a scenario approach to both requirements analysis (design) and assessment (evaluation) of the CoMMA platform – requirements analysis and evaluation being interleaved: "For requirements analysis, the aim is to 'get at' the user needs; for evaluation the aim is to 'tune' the system to make sure that it really does meet those needs" [25].

The requirements analysis of CoMMA was initially oriented by the two following scenarios: *(1) NEI Scenario:* The "integration of new employees" in a company; it concerns the new employees who need to handle a lot of new information about their enterprise in a very short time, to be rapidly efficient; *(b) TM Scenario:* the diffusion of innovative ideas among employees particularly when dealing with "technology monitoring activities;" it concerns the necessity for each enterprise to access in a very effective way to information concerning technology movement through the Internet that could contribute to its development. These scenarios were originally committed to the two industrial partners of the CoMMA project – a German and an Italian telecommunication company – who took the role of the application end-users.

Because the two scenarios were very abstract and vague, we needed to specify them to get requirements that could be converted into operational system specifications. How did we achieve this? We could have applied the "motivating scenario" approach, now classical within the ontology engineering community, and which underlies the TOVE ontological engineering method elaborated by Grüninger and colleagues [10][19][20].

> *Motivating scenarios* are a core component of the TOVE method. The notion of a motivating scenario refers to a "detailed narrative about the enterprise where emphasis is placed on the problems that the enterprise is facing or the tasks it needs to perform to solve the problems" (e.g., improving enterprise planning and scheduling). Ontology engineers use these problems to define an ontology's requirements in the form of *competency questions* that an ontology must be able to answer (e.g., *What sequence of activities must be completed to achieve some goal? At what times must these activities be initiated and ter-*

minated?). The competency of the ontology is tested by proving completeness theorems with respect to the competency questions.

We however found a main limitation to the "motivating scenarios approach:" the informality and user-orientation present in the first steps of the OBT design process were lacking in the evaluation step. We needed a more user-centered approach. Hence we turned towards the scenario approaches proposed by the Human-Computer Interaction (HCI) and Computer-Supported Cooperative Work (CSCW) communities, e.g., the approach of Carroll and his colleagues [4][5].

3 Exploiting Scenario Approaches Familiar to the HCI and CSCW communities

The scenario approaches have been introduced in the HCI and CSCW communities to fill the gap that the "traditional approach" to design created by imposing a technological orientation, abstraction, and other "user-distant" features. The scenario approaches allowed a design team to reintroduce the user's viewpoint in the design cycle, and to take into account her need of speaking of the system in terms of the work she has to achieve, using concrete and specific terms, and so on. In Carroll and colleagues approach, for example, scenarios of use are defined as descriptions, often narratives, of what people (could) do and experience (e.g., problems) when using computer systems. Scenarios can be developed through direct observation of users performing tasks in their work environment (observed scenarios), or through abstractions from theories of human activities (envisioned scenarios).

By exploiting the scenario approaches of the HCI and CSCW communities, our aim was to balance technology-orientation (prevalent in the ontology engineering community) with user-orientation (recommended by the HCI and CSCW communities), and, more specifically, to balance formality (which is a strong standard within the ontology engineering community) with informality (which is a HCI and CSCW requirement for not losing touch with the user, see e.g. [3]).

Scenarios are a meaningful way of accounting for users' needs. They embody properties, qualities or criteria that must be "put" in the system so that the system be accepted by its intended users. Scenarios embody criteria that must be found when assessing the system. Scenarios are both requirements and assessment scenarios. This two-sided aspect of scenarios would need to be systemized.

4 Eliciting Scenarios for Requirements Analysis

Applying a Scenario Approach Supposes To Have a Model of Scenario. The scenario formats an techniques proposed by the HCI and CSCW communities are multiple: e.g., scenarios, use cases, examples, stories, narrative descriptions of context, mock-ups, etc. We did not privilege a particular technique or format, but collected from the existing ones the elements that could help us answer methodological questions like: Which types of scenarios did we want to produce? Which contents shall we give to these scenarios? Which procedures are worth to follow to fill the scenario slots? As a result, in collaboration with the industrial partners, we elaborated a scenario grid to be used for requirements analysis by the partners (see Table 1).

Table 1. The CoMMA scenario elicitation grid

CHARACTERISTICS	REPRESENTATIONS	FACETS	
Goal: *Before:* *After:* *Scope:* *Scenario* *Sub-Scenario:* *Generic* *Specific* *Example* *Relevance life-time* *Exceptions* *Counter examples*	*Textual :* *Graphical :* *Informal :* *Formal :* (e.g., UML) In one scenario description, several types of representations may be used.	*Actors* Profile Role Individual goal Task Action Interaction *Logic & Chronology* Processes Decomposition Sequential/Parallel/ Non-deterministic Loops & Stop conditions Alternatives & Switches Compulsory/Optional *Functionalities & Rationale* Functionalities description Motivation, necessity Advantages & Disadvantages	*Resources* Nature Services Constraints *Flows* Inputs Outputs Paths *Environment* Internal Organization Acquaintance External

In the grid, "Characteristics" and "Representations" allow to specify the type of scenario to elicit, e.g., a scenario informally describing, in a textual format, a specific existing situation. "Facets" refer to the contents of the scenario: the actors involved in the scenario, having certain roles, using certain resources, performing a certain task in a certain way, and so on. To each of the elements of the scenario grid, we associated definitions and examples, and also questions to help industrial partners elicit relevant knowledge. For example, to the "Actors Interaction" facet, we associated the questions: *Who helps you to perform your job? Which persons do you consult to get information ... / to get that ... done?*

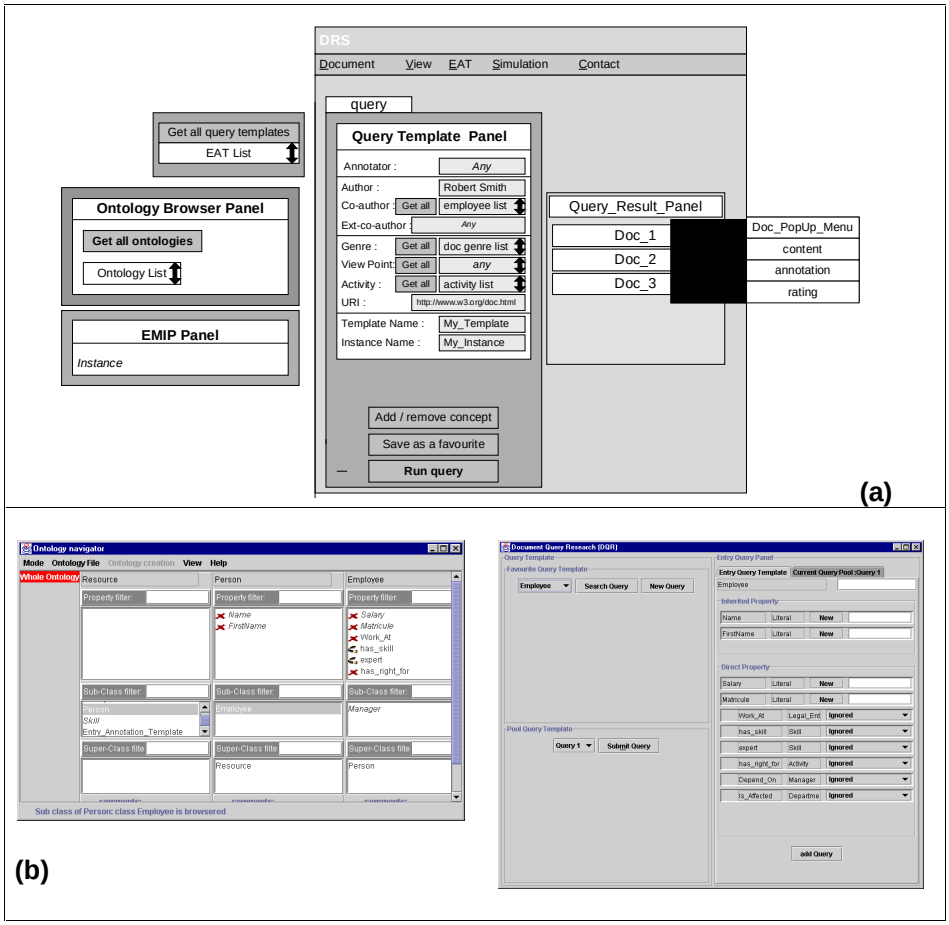

Fig. 1. The first CoMMA interfaces. Example of the Document Retrieval System (DRS) for the Technology Monitoring scenario: *(a)* as designed; *(b)* as implemented (*Left:* the "Ontology Browser Panel;" *Right:* the "Entry Query Template" corresponding to the selection of a concept into the Ontology Browser Panel)

To each of the scenario elements, we also associated techniques and potential corporate sources that can be exploited to answer the questions. The grid was given to the industrial partners who found them helpful for requirements analysis. A number of scenarios significant to end-users were thus elicited.

Applying a Scenario Approach Supposes a Continuous Focus on Scenarios. The CoMMA project was divided into two phases, each one ending with a trial at each industrial partner site. We must admit that, during the first phase, after requirements analysis, we lost sight of the scenarios, and that we consequently lost touch with the end-users. There were two reasons to this: (1) end-user partners were not truly available for the trial preparation and execution (one of them even withdrew from the Consortium, and was replaced by another partner belonging to another industrial sector: construction); (2) the priority of the research partners in the first phase was to perform and test the integration of new technologies, not really to meet end-users' needs. The result of this distance from users was foreseeable: immersed in technology, abstraction, and formality, the Consortium designed interfaces for developers and ontologists, and not interfaces for end-users. Figure 1 gives an idea of these interfaces for technologists. So it is not enough to have an early focus on scenarios of use and on users, it is necessary to have a continuous focus on them.

4 Using Scenarios for Tool Assessment

Using Scenarios to Assess Functionalities with Users. As a consequence of the distance from users, our interfaces were definitely not usable by end-users, and direct usability testing of these interfaces by end-users was impossible. Being however convinced that developing an OBT was a promising solution for supporting corporate memory management, we decided not to give up, and to show the interest of the CoMMA solution to potential end-users by making the CoMMA functionalities tangible through scenarios familiar to users. The goal was to describe the functionalities in terms of the work users will perform with the system.

We illustrated the functionalities of CoMMA through various scenario formats, in particular *storyboards* – a scenario format mixing text and images – of actual information-seeking newcomers' activities within the intranet of one end-users' company. These scenarios allowed to identify specific processes likely to be performed when using the system, e.g.:

> *"Travel Expenses Refund" scenario (excerpt).* A newcomer was seeking instructions in the intranet of his company for the refunding of his travel expenses. During the information-seeking process, the newcomer proposed different keywords to the successive search engines he utilized, or he followed various links related to "Travel expenses." Table 2 provides the sequence, and the transformation, of keywords entered and links followed by the newcomer during his activity. The contents of Table 2 illustrates a user's continuously changing process that we can call "term/concept shifting" (further discussed below).

Table 2. The series of transformations of the keywords used by a new employee searching for instructions for the refunding of his travel expenses

GERMAN KEYWORDS OR LINKS USED	ENGLISH TRANSLATION
Reisekostenabrechnung	**Travel expenses account**
Reisekosten**richtlinie**	Travel expenses **guideline**
R~~eisekostenrichtlinie~~ *(R refers to: Reisekosten~~richtlinie~~)*	Travel expenses ~~**guideline**~~
Reise~~**kostenantrag**~~	Travel ~~**expenses**~~ **request**
Reise~~**antrag**~~	Travel ~~**request**~~

(*Convention:* The part of the keyword which did not change from the previous turn to the current turn is printed in gray.)

Going back to the scenario approach, and consequently getting again in touch with end-users, we can show that end-users found the functionalities very useful, and suggest refinements and extensions to these functionalities (e.g., term/concept shifting illustrated in Table 2).

Using Scenarios to Assess Interfaces without Users. Using two scenario-based techniques – namely, Heuristic Evaluation [23] and Cognitive Walkthrough [27] –, we were all the same able to assess the CoMMA interfaces without users, but by putting ourselves in the users' shoes. This indirect assessment permitted us to identify usability problems, to propose recommendations for overcoming them, and to suggest interface specifications based on these recommendations (for details, see [15][16]). The two scenario-based methods have the following advantages:

Contextualizing Assessment Criteria through Scenarios. The Heuristic Evaluation technique consists for the evaluator in looking for violations of common usability principles or heuristics, such as *Flexibility and efficiency of use*: "Accelerators – unseen by the novice user – may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions." The usability inspection is greatly facilitated when evaluators are provided with scenarios: the criteria being affected by the context of use (i.e., user's characteristics, task, environment), scenarios allow to contextualize the criteria, and to make them meaningful. For example, "concept/term shifting" of "Travel Expenses Refund" scenario can be related to a Flexibility issue.

Justifying Scenarios with Activity Models. The Cognitive Walkthrough technique consists for the evaluator in "walking through" the interface, trying to act as a user. The walkthrough process involves examining each individual action step and trying "to tell a believable story" (scenario) about why the prospective user would choose an action. Scenarios in Cognitive Walkthrough are based on, and justified by, a model of exploratory learning of the system, which describes human-computer interaction in terms of four steps:

1. The user sets a goal to be achieved with the system (for example, “I am searching for corporate instructions for the refunding of my travel expenses”).
2. The user searches the interface for currently available actions (menu items, buttons, ontology browsing, etc.).
3. The user selects the action to progress toward the goal (e.g., browse the lists of concepts/terms for the concept/term “Travel expenses guideline”).
4. The user performs the selected action and evaluates the system's feedback for evidence of her progress (e.g., “I see that the term ‘guideline’ doesn't exist to refer to the concept of ‘instructions,’ but a synonymous term exists, that I can use to access to the corporate document I need”).

If we admit that the strength of an assessment method like the Cognitive Walkthrough depends on the relevance of its underlying model, a further step in the systemizing of the scenario approach would be to propose other models of human activity to justify the scenarios, e.g., models of the users’ linguistic activity. For example, the “Travel Expenses Refund” scenario could be explained by the notion of “concept drift” used in the Machine Learning community [22]. It can be also explained by the “vocabulary problem” model [11].

The models we spoke about so far are models of a user's individual activity, which are the most familiar to the HCI community. If we consider OBTs as tools supporting collective activities (e.g., elaborating a common terminology, sharing knowledge, etc.), we will need to refer also to models of collective activity, which are most familiar to the CSCW community. For example, the “Travel Expenses Refund” scenario could be justified by models like “lexical entrainment” [2], “concept and terminology co-ordination” [14], or “ontological drift” [24].

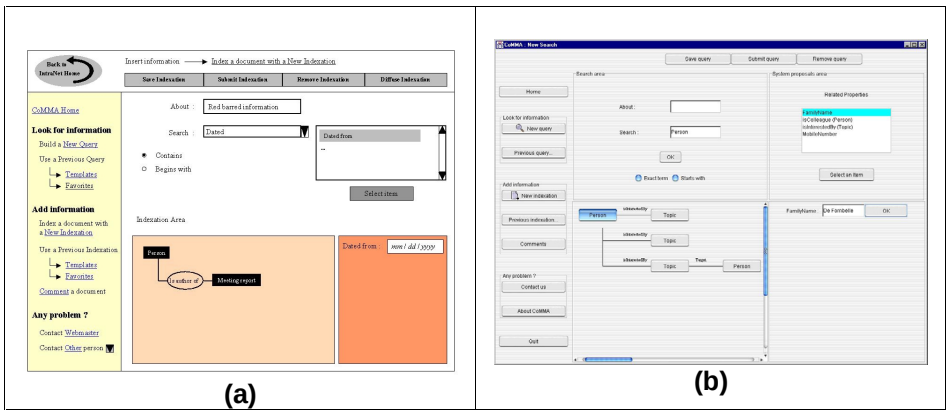

Fig. 2. The second CoMMA interfaces. Example of the Document Retrieval System (DRS) for the New Employee scenario: (a) as designed (with PowerPoint); (b) as implemented

Using Scenarios to Assess Interfaces With Users. For the second phase of the CoMMA project, we indeed made the necessary arrangements for not losing sight of the scenario approach, and not losing touch with the end-users. Among the arrangements we made were the following ones: (a) creating a HCI group, including, among others, end-users, interface developers, and human factors specialists; (b) involving the group in an iterative cooperative design/evaluation process; (c) inciting the group members to use scenario-based representations to discuss about the design and evaluation of the new interfaces. . As a result, we got simplified interfaces, that made sense to the users, and which users found this time usable (see Figure 2; for details of Trial 2, see [9]). However the process was very time-consuming.

5 Conclusion

In their *"Whitepaper: Evaluation of Ontology-based Tools,"* Angele and Sure [1] encourage the ontology engineering community, and more broadly the semantic web community, "to enforce their research efforts by developing further standard criteria [...] and tools that implement these criteria to evaluate ontologies and related technologies." Through the present paper, we tried to contribute to these efforts, showing for example that criteria development cannot be considered in isolation from situations in which the ontology-based tools will be used: to be meaningful and relevant, criteria need to be connected to scenarios of use, and these scenarios to be explained and further analyzed need to be connected to activity models. Put in other words, we claimed in this paper for a balance between usage and technology, and between formality and informality; in fact we advocated for avoiding premature formalization (as pointed out by Buckingham Shum [3]), or reinstalling informality when interacting with end-users.

Through this paper we invite the community to bring some efforts to bear on systemizing the scenario approach to assessment (and design), an approach more developed in the HCI and CSCW communities than in the ontology engineering community. It would be desirable to discuss also how to systemize the scenario approach for technical evaluation; the work by Kazman and his colleagues [21] on scenario-based evaluation of architectures is worth considering in such a discussion.

Acknowledgements. We would like to thank our colleagues of the CoMMA Consortium for their involvement in the assessment of the COMMA platform, especially: Sabine Delaître (CSTB), Bruno Fiès (CSTB), Karim de Fombelle (Atos Integration), Joachim Hackstein (T-Systems Nova), Hervé Karp (Atos Integration), Jérôme Leytier (INRIA), Emmanuelle Loyson (CSTB), Philippe Pérez (Atos), and Marie Soleilhet (Atos Integration). We would also like to thank two anonymous reviewers for their comments on a previous version of this paper.

References

1. Angele, J., Sure, Y.: *Whitepaper: Evaluation of Ontology-based Tools*. Excerpt from the IST-2001-29243 Report, OntoWeb. D1.3. Tools. (2001). Available at: http://www.aifb.uni-karlsruhe.de/WBS/ysu/publications/eon2002_whitepaper.pdf
2. Brennan, S. E.: Lexical entrainment in spontaneous dialog. *Proceedings of the 1996 International Symposium on Spoken Dialogue*, Philadelphia, PA: ISSD-96 (1996) 41-44. Available at : http://www.psy.sunysb.edu/sbrennan/papers/brenISSD.pdf
3. Buckingham Shum, S.: Balancing Formality with Informality: User-Centred Requirements for Knowledge Management Technologies. *AAAI Spring Symposium on Artificial Intelligence in Knowledge Management* (1997), Stanford University, Palo Alto, CA. AAAI Press. Available at: http://kmi.open.ac.uk/people/sbs/org-knowledge/aikm97/sbs-paper1.html
4. Carroll, J.M.: *Making Use: Scenario-Based Design of Human-Computer Interactions*. MIT Press, Cambridge, MA (2000)
5. Carroll, J.M., Mack, R.L., Robertson, S.P. & Rosson, M.B.: Binding objects to scenarios of use. *International Journal of Human-Computer Studies 41* (1994) 243-276.
6. CoMMA Consortium: CoMMA: Corporate memory through agents, *Proceedings of E-Work and E-Business'2000* (2000)
7. Corby, O., Dieng, R., Hébert, C.: A Conceptual Graph Model for W3C Resource Description Framework, *Proceedings of ICCS 2000*, Darmstadt, Germany (2000)
8. Corby, O., Faron-Zucker, C.: Corese: A corporate Semantic Web engine. *Proceedings of the International Workshop on "Real World RDF and Semantic Web Applications,"* WWW'2002, Hawai. (2002)
9. Fiès, B., (Ed.) (2002). Assessment Report of CoMMA Trial-step 2, CoMMA project Deliverable.
10. Fox, M.S., Grüninger, M..: Enterprise Modelling, *AI Magazine* (1998) 109-121.
11. Furnas, G.W., Landauer, T.K., Gomez, L.M., and S.T. Dumais.: The Vocabulary Problem in Human-System Communication, *Communications of the ACM 30* (1987) 964-971.
12. Gandon, F.: *Ontology Engineering: A Survey and a Return on Experience.* INRIA Research Report # RR4396, INRIA, France (2002). Available at: http://www.inria.fr/rrrt/rr-4396.html
13. Gandon F., Dieng R., Corby O. et Giboin A.: A multi-agent system to support exploiting an XML-based corporate memory, *Proceedings of PAKM'2000, the Third International Conference on Practical Aspects of Knowledge Management*, Basel, Switzerland (2000)
14. Garrod, S. How groups coordinate their concepts and terminology: implications for medical informatics. *Proceedings of the WG6 IMIA Symposium on Concepts and Terminology*, Jacksonville, Fl. (1997) 279-284
15. Giboin, A, Pérez, Ph. (Eds.): *Assessment Report of [CoMMA] Trial-step 1. Part 1 Technical Evaluation*, CoMMA Project (IST-1999-12217) Deliverable # COMMA/WP6/D10, 69 pages (2001a)
16. Giboin, A, Pérez, Ph. (Eds): *Assessment Report of [CoMMA] Trial-step 1. Part 2 User Evaluation*, CoMMA Project (IST-1999-12217) Deliverable # COMMA/WP6/D10, 156 pages, (2001b)
17. Gómez-Pérez, A.: *Some Ideas and Examples to Evaluate Ontologies*. Technical Report # KSL-94-65, Knowledge Systems Laboratory. Stanford University (1994a). Available at: http://www-ksl.stanford.edu/KSL_Abstracts/KSL-94-65.html

18. Gómez-Pérez, A.: *From Knowledge Based Systems to Knowledge Sharing Technology: Evaluation and Assessment*. Knowledge Systems Laboratory, Technical Report # KSL-94-73 (1994b)
19. Grüninger, M., and Fox, M.S.: The Role of Competency Questions in Enterprise Engineering, *Proceedings of the IFIP WG5.7 Workshop on Benchmarking – Theory and Practice*, Trondheim, Norway (1994)
20. Grüninger, M., and Fox, M.S.: Methodology for the design and evaluation of ontologies, *Proceedings of the IJCAI Workshop on Basic Ontological Issues in Knowledge Sharing*, AAAI Press, Menlo Park CA, (1995). Available at: http://www.ie.utoronto.ca/EIL/public/org.ps
21 Kazman, R., Carriere, S. J., Woods, S. G.: Toward a discipline of scenario-based architectural engineering, *Annals of Software Engineering 9* (2000) 5-33.
22. Lane, T. and Brodley, C.E.: Approaches to online learning and concept drift for user identification in computer security. *Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining* (1998) 259-263.
23. Nielsen, J.: Heuristic evaluation. In: Nielsen, J., and Mack, R.L. (Eds.), *Usability Inspection Methods*, John Wiley & Sons, New York, NY (1994)
24. Robinson, M. and L. Bannon: Questioning representations. *Proceedings of ECSCW'91, the Second European Conference on Computer-Supported Cooperative Work*, Amsterdam, The Netherlands (1991). Available at: http://www.ul.ie/~idc/library/papersreports/LiamBannon/15/QuestFin.html
25. Thomas, P.J. (Ed.): *CSCW Requirements and Evaluation*, Springer Verlag, London (1996)
26. Uschold, M. and Grüninger M.: Ontologies: Principles, methods and applications. *Knowledge Engineering Review 11* (1996). Also available as AIAI-TR-191 from AIAI, The University of Edinburgh.
27. Wharton, C., Rieman, J., Lewis, C., and Polson, P.: The Cognitive Walkthrough method: A practitioner's guide. In J. Nielsen and R.L. Mack (Eds.), *Usability Inspection Methods*, New York: John Wiley & Sons, (1994) 105-141.