The Value Grid for Semantic Technologies

Alistair Miles

STFC e-Science Centre, STFC Rutherford Appleton Laboratory, Chilton,
OX11 0QX, UK
a.j.miles@rl.ac.uk

Abstract

This paper situates formal ontologies as one of many products in a multi-tier *value grid* of semantic technologies. Incremental strategies for the exploitation of intermediate products in the value grid are discussed, as a possible step towards cost-effective, low-risk and scalable business models for the exploitation of semantic technologies. A case study is presented, illustrating a hypothetical value grid for the management of scientific data from a large-scale experimental facility. Suggestions are made for the design of predictable, repeatable collaborative processes for adding value in semantic technology value grids.

1 Introduction

An ontology is the product of a group of people, collaborating to articulate their commonly held conceptualisation of a domain, through the use of a formal logical language such as KIF or OWL [1, 2]. Typically, an ontology is intended for deployment in software systems which leverage the shared conceptualisation to provide unique, high-value services, such as data integration or information retrieval.

Ontologies are not the only means of articulating a shared conceptualisation, however. Controlled vocabularies, taxonomies, thesauri, classification schemes, topic maps, subject heading systems, semantic networks – to name a selection – are all specifications of a shared conceptualisation, albeit “informal” or “semi-formal”. These and many other types of product have to be considered, in order to design solutions to specific problems at reasonable cost; solutions that are feasible, scalable and part of a sustainable business model. This begs a number of questions. What possible paths exist from knowledge expressed informally (unstructured information) to formal ontologies? How can these paths be broken down into stages, and what does each stage produce? In what ways can these different products be exploited? What are the likely costs, benefits and risks associated with different paths and different stages? How much human effort will be required, and how can this effort be reduced by computation? How can the necessary human effort be organised into efficient work flows that enable collaboration? Does economic and practical scalability vary with different paths and different products?

Value chains and value grids

This paper works towards answers to these questions, by viewing ontologies as products in a *value grid* of semantic technologies. The notion of a “value grid” has evolved from the original conception of a “value chain”, which is a sequence of value-enhancing activities, where raw materials are formed into components that are assembled into final products, distributed, sold and serviced [3]. The “value grid” extends this view, to see value creation as multi-dimensional rather than linear. In a value grid, the *vertical* dimension describes multiple *tiers* from primary inputs (raw materials) to end users; the *horizontal* dimension describes opportunities at the same tier across parallel value chains; and the *diagonal* dimension describes opportunities for integration between value chains [3]. *Value grids*, rather than *value chains*, are used here, because the extra degrees of freedom allow for a subtle analysis of relationships between products such as thesauri and ontologies, where several possible configurations can be discussed.

By exploring and “mapping-out” the value grid in which ontologies are situated, we may begin to define profit-maximising strategies for the exploitation of semantic technologies. Such strategies might, for example, identify and incrementally exploit a fine-grained sequence of products leading from unstructured information to formal ontologies. By taking this incremental approach, a project would not “overshoot” its requirements and formalise beyond what is necessary or worthwhile. A project would also be able to return value to its customers early and often. If there is a tight cycle whereby stakeholders perceive early returns on small levels of investment, they may be inspired to deepen their initial commitment.

A heavyweight "all-or-nothing" approach to ontology engineering demands high levels of stakeholder commitment, both initially and on an ongoing basis, with delayed returns. Frustration can ensue because stakeholders cannot be persuaded to make the commitment required by such an approach, nor can they be persuaded to take ownership of the product. In some situations a formal ontology is the only product that can support the desired functionality, and therefore an "all-or-nothing" approach is appropriate. However, there are many other situations in which other products in the value grid may be exploited, and therefore other strategies become available.

Section 2 begins an exploration of the value grid by re-analysing a case study involving the management of scientific data from a large scale facility. The case study explores options for adding value to a catalogue of experimental data, for which uncontrolled keywords are already present.

Collaboration engineering

Almost all value-adding activities in the exploitation of semantic technologies require at least some human intellectual input. Moreover, because semantic technologies demand *shared* conceptualisations, these activities are necessarily collaborative. This dependence on collaboration is a major source of both cost and risk *throughout* the value grid, because designing and managing predictable, repeatable collaborative processes is a significant challenge.

The successful execution of a collaborative process traditionally depends on the involvement of a professional facilitator – someone who can design and enact a dynamic process, structuring tasks and managing relationships between people, tasks and technologies [4]. However, the continuous involvement of a professional facilitator is expensive. To reduce this dependence on professional facilitators, the field of *collaboration engineering* has sought to *codify* and *package* key "facilitation interventions" in forms that can be executed by team members themselves [4]. A "collaboration engineer" designs a group process in a way that can be transferred to a "practitioner" – a domain expert who can execute a single team process as a team leader in their particular domain. The collaboration process is broken down into a set of atomic collaborative activities ("thinklets"). Each of these basic units comprises a named, packaged, thinking activity that creates a predictable, repeatable pattern of collaboration among people working towards a goal. I.e. each unit provides a concrete group-dynamics intervention, complete with instructions for implementation as part of some group process [4].

Collaborative activities may be characterised in a variety of ways [4]. For example, an activity can be classified according to the pattern of collaboration: *divergent* activities move from fewer to more concepts; *convergent* activities focus attention from many to fewer concepts; *organising* activities increase understanding of relationships between concepts; *evaluating* activities assess concepts relative to some criteria. An activity can also be classified according to its outcome: whether the product is an unstructured collection of concepts; an overview; a structure such as a list, tree or directed graph; and whether the output has been "judged" and/or "cleaned" [4]. These and other tools from collaboration engineering are used in section 3 to analyse traditional assumptions about collaborative ontology engineering, and to begin a detailed analysis of collaborative activities involved in the construction of a shared conceptualisation.

Section 4 discusses a selection of relevant work, and section 5 presents conclusions and further work.

2 Value Grid Case Study

This section describes a case study, illustrating a possible value grid for semantic technologies to improve the management of experimental data from the ISIS facility [5]. Currently, metadata describing experimental outputs is managed via a metadata catalogue in which one or more uncontrolled keywords may be associated with each experiment. However, many of the keywords are synonyms, and the keywords constitute a local dialect that can be hard to penetrate for users of other, similar facilities [6] – therefore current retrieval services are limited.

This case study was the subject of a recent project, exploring the feasibility of using a formal ontology for enhanced retrieval services [6]. An ontology was developed to replace the keyword indexing system, in collaboration with

ISIS scientists. However, a number of difficulties were encountered, not least the lack of consensus on a best-practice methodology for ontology engineering, and the difficulty of facilitating collaboration between ontology engineers and domain experts – in part due to the inability of domain experts to comprehend and use ontology engineering tools [6].

What products, other than an ontology, might be exploited to improve retrieval services for ISIS experimental data, and how might these be developed and deployed as part of an incremental strategy?

If an incremental strategy is followed, all possible value will be extracted from currently held assets, before any investment of effort is made in the development of further products. An uncontrolled keyword vocabulary is in itself an important product in the value grid. To maximise the value being obtained from this asset, the keyword index could be exploited to provide a number of additional services, both to users searching for experimental data (*searchers*) and to users entering a description of their own experiments (*submitters*).

For the *searcher*, the process of entering and refining/modifying search queries could be supported in several ways. Query hints and suggestions could be provided, which could be either passive or active. A passive example would be to present a visualisation of the keyword index to the searcher (e.g. as a "tag cloud"). An active example would be to perform sub-string matching on queries as they are being typed, and suggest completions from the current keyword vocabulary. The keyword index could also be used as the subject of a statistical analysis to crudely identify clusters of related experiments. Clusters could be used to group results within large result sets, and to provide "see also" links between individual results.

For the *submitter*, an "auto-completion" suggestion feature could be provided when entering keywords. Keywords could also be suggested prior to entry, based on a direct analysis of the text of the experiment title and abstract, or perhaps based on a more sophisticated statistical comparison of the title and abstract with other titles and abstracts already indexed with keywords.

These additional services might also lead to a marginal increase in the *quality* of the keyword index, by providing paths for feedback between keyword entry and keyword use. If users are able to perceive the impact of their keyword assignments and those of their peers, and can adapt their usage depending on the behaviour of others in the community, this could, at least in principle, lead to communities and patterns of use emerging – this is the theoretical basis for "social tagging".

A second product in the value grid might be a set of synonym (equivalence) links between currently used keywords. Synonym links could be immediately exploited in a number of ways. For the *searcher*, synonyms could be used to provide suggestions for alternative queries. Synonyms could also be used "behind the scenes" to expand queries, increasing recall without demanding any additional action from the searcher. Synonyms could also improve the performance of natural language processing techniques, because they remove the necessity to perform co-reference resolution [7]. This in turn could improve the analysis of available text in titles and abstracts, leading to better suggestions for keyword entry and better clustering of related experiments.

If a preference is expressed for one keyword in each synonym set, then a primitive controlled vocabulary is produced. This preference could be indicated to the *submitter* during keyword entry, which could influence convergence in keyword usage, without restricting the freedom of the keyword system. Note that vocabulary control is also a prerequisite for most products in the value grid involving formalisation of syntax and/or semantics.

Once keyword synonym sets have been identified, various products can be created which organise these sets in different ways.

One such product involves a high level categorisation, such that all synonym sets are placed into one of a small number of categories (a.k.a. "facets"). Casely-Hayford & Sufi found several high-level categories, including the experimental instrument, the subject of the investigation, the investigating body/group and the year of the experiment [6]. By producing a high-level categorisation of all synonym sets, various options immediately become available for provision of services to both searcher and submitter. For the *searcher*, "faceted" search/browse interfaces can be constructed, allowing users to build composite queries involving multiple categories, such as searching for experiments on a particular instrument by a particular group in a particular year. For the *submitter*, the submission form could be structured, and a smaller number of suggestions could be given for keyword values specific to each category.

Another, complementary, option would be to organise synonym sets into hierarchies (trees) and/or to find associative links between sets. Hierarchies can be exploited in a number of ways. For example, suggestions could be provided to the *searcher* for making their current query either more general or more specific. Hierarchical relationships could also be used behind the scenes to expand keyword queries, further improving recall. Another possibility is to offer hierarchies as a means of browsing a set of results. Associative links could be exploited to provide additional "see also" links for browsing result sets.

Another value-adding activity would be to annotate synonym sets with small amounts of explanatory text. An annotated vocabulary could be exploited in a non-intrusive way to assist both searcher and submitter, e.g. by using annotations as the content of "tool tips" associated with keywords displayed in user interfaces.

In sum, there are many ways in which value could be added to the current keyword system, without going as far as the development of a formal ontology. There are also ways of adding value, without enforcing strict vocabulary control. A range of products can be identified, including an uncontrolled keyword vocabulary; synonym sets; primitive controlled vocabulary; categorised (faceted) vocabulary; structured vocabulary (primitive thesaurus); and annotated vocabulary. Each of these products could be exploited to provide new features. Each of these products also, to a certain extent, could provide input ("raw materials") to the development of other products, including those at higher levels of formalisation.

A viable incremental strategy might be to develop and exploit products in the order introduced above.

3 Collaboration

All semantic technologies work from the assumption that a *shared conceptualisation* is captured in some sort of *specification*, from which various useful and unique functionalities are then derived. Under this assumption, an information system will only be useful to those people who actually share the conceptualisation which is deployed therein. Of course, a single person could be employed to attempt to capture, integrate and articulate a conceptualisation held by others. However, it is generally assumed necessary for knowledgeable members of the application domain ("domain experts") to be involved in the conceptualisation process, so that their views may be represented directly. To achieve some assurance of "sharedness", more than one person must be involved, and therefore the conceptualisation process demands collaboration. Effective collaboration is a critical factor in the successful application of any semantic technology.

This section examines the nature of collaboration in the application of semantic technologies, beginning with assumptions about collaboration in ontology engineering.

Collaborative ontology engineering

Methodologies for ontology engineering typically assume that participants in the process play one of two roles: either *ontology engineer* or *domain expert* (see e.g. [8]). The conceptualisation which is shared by the domain experts is to be captured in and expressed by the ontology. The ontology engineer is responsible for implementation of the ontology in an ontology language. Beyond these two statements, little is said about how the domain expert and the ontology engineer should interact during the conceptualisation process, or indeed what their specific goals are. Is an ontology engineer supposed to translate informal statements made by the domain experts into formal statements in terms of an ontology language? Is the ontology engineer supposed to educate the domain experts in ontology modelling, and encourage them to structure their thoughts in terms of classes, properties, individuals, axioms etc.? Should the ontology engineer broker agreements between domain experts in an attempt to reach consensus? Is it realistic to expect domain experts to carry out discussions using a formal ontology language? Three distinct collaborative relationships can be identified: relationships between domain experts; relationships between ontology engineers; and relationships between domain experts and ontology engineers. What are the natures and goals of these different collaborations? How may they be facilitated and supported by software tools?

Many, if not all, of the currently available tools intended to support "collaborative" ontology development are unclear as to *who* they are intended to support collaboration *between* (although cf. [9]). Some projects have found that domain experts lack both the experience and the willingness to engage with ontology development tools. This means that, if a commitment is made to formal ontology engineering, other tools have to be built in order to enable communication between ontology engineers and domain experts (e.g. the "OntoMaintainer" [6]). What of tools to enable direct collaboration between domain experts? Without these, how are domain experts supposed to "share" their conceptualisation?

As implied in the previous section, some value chains – paths through the value grid – might employ logical formalism only during the later stages of production, if at all. Activities such as finding synonym links between keywords, organising synonym sets into categories, hierarchies and networks, may be quite intuitive, and certainly won't require any experience of ontology engineering. Who, then, should be involved during different stages of the conceptualisation process? What is their role, and what skills do they require?

Conceptualisation processes

The study of collaboration must be the first consideration in the design of new software tools to support the conceptualisation process; tools specifically designed to support *direct facilitated collaboration between domain experts*. It is beyond the scope of the current paper to undertake such a study in any depth. However, below families of collaborative activities are sketched in outline; activities which could form the basis for the design of predictable, repeatable processes for the construction of a shared conceptualisation. The aim is to suggest ways of breaking the conceptualisation process down into a set of composable activities – building blocks which could be arranged into collaborative processes *ad hoc* in order to construct different products and traverse different paths in the semantic technology value grid. Possibilities for supporting these activities with "computer aid" are also discussed.

In the case study given above, an uncontrolled keyword vocabulary was already present. However, many projects will have to start from scratch. Therefore, a first family of collaborative activities are those directed towards the collection of "raw materials" for the conceptualisation process – objects which convey informal expressions of meaning, without any context or structure, such as keywords, fragments of text, images, audio or video clips etc. All activities in this family are divergent – the aim is to obtain a comprehensive collection of everything that *might* be relevant.

One concrete example of an activity in this family is "word association" – members of a group are asked to propose words or phrases in association with prompted suggestions. The group could continuously prompt each other, or could be prompted from a number of predefined starting points. A second concrete example is "literature scanning", where individuals read documents and extract important words or phrases [10]. Both of these activities benefit from rich interaction between participants, and may be "computer aided" e.g. via statistical analyses of textual material. Another, quite different, concrete example of an activity in this family is "social tagging", where individuals use their own keywords to describe objects of interest, generating an "folksonomy".

In the ISIS case study, the first proposed activity was the establishment of synonym links between keywords. We can generalise this to identify a family of activities, whose goal is to organise objects collected as "raw materials" into "synonym sets", where each set of objects provides at least a partial indication or perception of a distinct "concept" to one or more persons in the collaboration (although "synonym" is used loosely here, especially if "synonyms" can include multimedia objects). Note that this activity is also divergent – the goal is to find all reasonable sets for all members of the collaboration, so that all views are initially represented. Variant concrete activities within this family might involve people working on an individual "work space", seeing other colleagues' sets only when made, or might involve all members of the collaboration working simultaneously in a shared work space, seeing and influencing each others' actions in real time.

Because the number of collected objects may initially be large, this second family of activities is also a candidate for "computer aid". Synonyms might be suggested via background sources, such as general thesauri or word nets; or from mathematical analysis of usage graphs in social tagging networks [11].

Another family of activities involves *judging* synonym sets – rating and voting on sets so that the "best" are kept and the "worst" are discarded. Sets are "better" if they provide a clear indication of a distinct "concept" that is recognised to some extent across the collaboration. This is a convergent activity, resulting in a "cleaner" collection [4]. However, activities in this family do not have to have absolute agreement – the aim could simply be to provide a candidate collection, deemed worthy of further evaluation and elaboration.

Once candidate synonym sets have been judged, for those that remain it is likely that further modifications will need to be made. Two families of activities for adding/removing objects from synonym sets can be identified: those making proposed changes (divergent) and those judging/voting and choosing alternative propositions (convergent). Adding textual annotations – complete or partial definitions – to synonym sets could be identified as a separate activity, carried out before and/or after judging of synonym sets. Effort could be prioritised in a number of ways, e.g. by targeting those accepted candidates for which there was the least consensus during judging. Adding annotations might also be carried out as part of judging, where individuals add short annotations in order to clarify meaning and "improve" or "promote" a set.

Several families of activities for organising synonym sets into higher-level structures can be identified, according to structures being produced. Structures include high-level grouping/categorisation; hierarchies (trees); association networks (graphs). Activities can also be characterised according to divergent or convergent aims; for example, some activities will collect proposals for structuring in different ways, whereas others will evaluate and choose between alternatives.

Of course, the process will have to reach a point where the conceptualisation is deemed "complete", i.e. sufficient to generate the desired product and achieve chosen quality criteria. This suggests various activities for the evaluation of a conceptualisation as a whole, and for collaborative decisions to "publish" (i.e. to release a new "edition").

4 Discussion

Thesauri and other types of KOS

The study of "knowledge organisation systems" (KOS) is highly relevant to an exploration of products in the semantic technology value grid. There are many different types of KOS, including thesauri, taxonomies, classification schemes and subject heading systems (although the distinctions are sometimes blurred, see [12] and [13]). Nevertheless, KOS generally provide a controlled vocabulary, may provide synonym links, and may organise their conceptual units into hierarchies and/or networks of association. Ontologies are sometimes viewed as a type of KOS, although ontologies are fundamentally different due to their formal semantics.

It is possible to use thesauri as input to a formal ontology engineering process (see e.g. [14] and [15]). However, a thesaurus or taxonomy is not necessarily an appropriate precursor to the development of a formal ontology. A thesaurus developed for one application (e.g. information retrieval) might be quite useless as input to the development of an ontology for another (e.g. database schema integration). Nevertheless, *activities* and *techniques* used in the development of thesauri, taxonomies etc. *might* be applicable to early stages in the development of a shared conceptualisation, depending on the ways in which products of the conceptualisation process are to be exploited.

Aitchison et al. [10] divide the process of constructing a thesaurus into two major phases: *term selection* and *finding structure*. Techniques for the collection of terms include selection from existing terminological sources (other thesauri, classification schemes, glossaries etc.), and manual scanning and/or automatic analysis of relevant literature. Finding structure begins with a preliminary organisation of terms into a few broad subject categories (e.g. "catering"). Within each category, basic facets are then recognised and stated (e.g. "equipment", "operations"). Within basic facets, terms are then arranged into hierarchies. Finally, scope notes, equivalence relationships, associative relationships and additional (poly-)hierarchical relationships are added. In the case study described in section 2, it was suggested that finding synonym links between keywords be done prior to finding other structures, because the synonym links can be immediately exploited for query expansion, natural language processing and other tasks. In contrast, Aitchison et al. find equivalence relationships *after* finding hierarchical and categorical structures.

In 10 two roles are identified: "thesaurus compiler" and "subject specialist" (cf. "ontology engineer" and "domain expert"). Although it is recommended these two roles should interact during construction of a thesaurus, it is clear that most of the actual work of selecting terms and, in particular, finding structures, is to be done by the thesaurus compiler. Little, if any, attention is paid to different ways of structuring collaborations within and between subject specialists and thesaurus compilers, in order to produce successful, repeatable, outcomes. This suggests that collaboration engineering has much to offer the study of collaborative processes in the construction of both thesauri and formal ontologies.

Statistical techniques and automation

Taxonomies are still a major topic in enterprise search, despite the movement of Web search engine providers into this space. There has been a dichotomy between advocates of solutions to enterprise information management based on librarianship, through vocabulary control and metadata, and those who argue that the traditional library business model does not scale with the volumes of information currently managed within a typical enterprise, and that therefore the only viable approach is via complete automation (see e.g. [16]) – which in turn depends on a variety of statistical techniques. This paper argues for a tight coupling of semantic technologies and technologies based on various techniques for the statistical analysis of text and other types of unstructured content. Such a coupling can be realised through an incremental approach to the exploitation of semantic technologies, because at each stage the products of the conceptualisation are used to improve the performance of statistical techniques, and statistical analyses are used to provide "computer aid" in the construction of the conceptualisation. In this way, the limited amount of human effort available may be used to the greatest possible effect, within a scalable business model.

All work on the analysis of unstructured information, including natural language processing, is therefore relevant. This is, however, a broad field, and a complete review of relevant techniques is well beyond this scope of this paper. Three references are selected for illustration. Welty & Murdock [17] perform knowledge acquisition from unstructured text, integrating components from the UIMA framework with Semantic Web tools. Ramakrishnan et al. [7] use controlled vocabularies and natural language processing to support automatic relationship discovery between unstructured texts. Mika [11] develops a model for the analysis of social tagging networks, to derive emergent structures, including crude synonym links and hierarchies.

Collaboration and Ontology Engineering

High-level patterns of collaboration in the construction of ontologies are characterised in [18], which also presents plug-ins to the Protege platform to support collaboration and change management. However, different roles in the collaborative process are not discussed. Sure et al. [9] go further, describing a plug-in for the OntoEdit platform specifically designed to support collaboration between an ontology engineer and a domain expert ("OntoKick"), and a plug-in to support the integration of mind maps (presumably the result of collaboration between domain experts) into the ontology engineering process ("Mind2Onto"). Neither study provides a detailed analysis of the process of conceptualisation, or any suggestions for individual collaborative activities to bootstrap the conceptualisation process. A collaborative approach to ontology design is proposed in [19], based on an adaptation of the Delphi method, a formal technique for the integration of multiple viewpoints. An initial "seed ontology" is developed, then iteratively improved in response to feedback elicited from a "panel" of domain experts via questionnaires. Again, however, domain experts are not given the opportunity to collaborate directly, only indirectly via the medium of an ontology engineer.

Recently, the "wiki" paradigm has been employed in the design of ontology engineering tools. Semantic MediaWiki [20] allows formal assertions to be included in wiki page content by means of an extension to the wiki syntax (assertions are interpreted as ABox statements in the OWL DL language). OntoWiki [21] takes a different approach, providing a collaborative editing environment based on customisable widgets, rather than a wiki syntax. Although both approaches are advertised as providing an intuitive interface to the ontology engineering process, both rely on a basic understanding of the underlying ontological primitives (classes, individuals, properties etc.). No support is provided for more informal techniques for developing a conceptualisation, such as those suggested in sections 2 and 3 above.

5 Conclusions

This paper has begun an exploration of fine-grained value chains in the application of semantic technologies. An incremental strategy is proposed, whereby all available products in a value chain are exploited in turn, and the greatest possible value is obtained from each subsequent product before a commitment is made to production of the next. In this way, customers receive a return on their investment early and often, and appropriate levels of investment can be found.

Collaboration is, by definition, an essential element in the social construction of a shared conceptualisation. This paper has begun an analysis of patterns of collaboration and collaborative activities that can be employed in the design of predictable, repeatable processes. The underlying goal of this paper is to work towards the design of software environments to support the collaborative construction of shared conceptualisation; designs which take as their starting point an understanding of the social activities which are to be enabled through the provision of shared, virtual, work spaces. A next step is therefore a detailed analysis of collaborative activities and the design of collaboration processes, in conjunction with the design of virtual work spaces and group support systems to support those processes.

References

1. Thomas R. Gruber, Toward Principles for the Design of Ontologies Used for Knowledge Sharing, 1993
2. Michael K. Smith, Chris Welty and Deborah L. McGuinness, OWL Web Ontology Language Guide, 2004
3. Frits K. Pil and Matthias Holweg, Evolving from Value Chain to Value Grid, Sloan Management Review (2006)
4. Gwendolyn L. Kolfschoten, Robert O. Briggs, Jaco H. Appelman and Gert-Jan d, ThinkLets as Building Blocks for Collaboration Processes: A Further Conceptualization (2004)
5. Damian Flannery, ISIS Data, Metadata, Proposals and the CCLRC Data Portal, Proc. NOBUGS Conference (2004)
6. Louisa Casely-Hayford and Shoaib Sufi, The ISIS Facilities Ontology and OntoMaintainer, Proc. UK e-Science All Hands Meeting (2006)
7. Cartic Ramakrishnan, Krys J. Kochut and Amit P. Sheth, A Framework for Schema-Driven Relationship Discovery from Unstructured text, Proc. 5th International Semantic Web Conference (2006)
8. Mariano Fernandez-Lopez and Asuncion Gomez-Perez, Overview and analysis of methodologies for building ontologies, The Knowledge Engineering Review (2002)
9. York Sure, Michael Erdmann, Juergen Angele, Steffen Staab, Rudi Studer and, OntoEdit: Collaborative Ontology Development for the Semantic Web, Proc. 1st International Semantic Web Conference (2002)
10. Jean Aitchison, Alan Gilchrist and David Bawden, Thesaurus construction and use: a practical manual (2000)
11. Peter Mika, Ontologies Are Us: A Unified Model of Social Networks and Semantics, Proc. 4th International Semantic Web Conference (2005)
12. Structured vocabularies for information retrieval - Guide - Part 2: Thesauri {2005)
13. Structured Vocabularies for Information Retrieval - Guide. Part 3. Vocabularies other than thesauri (2006)
14. B. J. Wielinga, A. Th. Schreiber, J. Wielemaker, J. A. C. Sandberg, From Thesaurus to Ontology, Proc. 4th International Conference on Knowledge Capture (K-CAP) (2001)
15. Dagobert Soergel, Boris Lauser, Anita Liang, Frehiwot Fisseha, Johannes Kei, Reengineering Thesauri for New Applications: the AGROVOC Example, Journal of Digital Information (2004)
16. Tom Reamy, To Metadata or not to Metadata, EContentMag.com (2004)
17. Chris Welty and J. William Murdock, Towards Knowledge Acquisition from Information Extraction, Proc. 5th International Semantic Web Conference (2006)
18. Natalya F. Noy, Abhita Chugh, William Liu, Mark A. Musen, A Framework for Ontology Evolution in Collaborative Environments, Proc. 5th International Semantic Web Conference (2006)
19. Clyde W. Holsapple and K. D. Joshi, A Collaborative Approach to Ontology Design, Communications of the ACM (2002)
20. Markus Krötzsch, Denny Vrandecic and Max Völkel, Semantic MediaWiki, Proc. 5th International Semantic Web Conference - Semantic Web Challenge (2006)
21. Soren Auer, Sebastian Dietzold, and Thomas Riechert, OntoWiki - A Tool for Social, Semantic Collaboration, Proc. 5th International Semantic Web Conference (2006)