Document Space Adapted Ontology: Application in Query Enrichment

Stein L. Tomassen, Jon Atle Gulla, and Darijus Strasunskas

Department of Computer and Information Science,
Norwegian University of Technology and Science,
Sem Saelandsvei 7-9, NO-7491 Trondheim, Norway
{stein.l.tomassen, jon.atle.gulla, darijus.strasunskas}@idi.ntnu.no

Abstract. Retrieval of correct and precise information at the right time is essential in knowledge intensive tasks requiring quick decision-making. In this paper, we propose a method for utilizing ontologies to enhance the quality of information retrieval (IR) by query enrichment. We explain how a retrieval system can be tuned by adapting ontologies to provide both an in-depth understanding of the user's needs as well as an easy integration with standard vector-space retrieval systems. The ontology concepts are adapted to the domain terminology by computing a feature vector for each concept. The feature vector is used to enrich a provided query. The ontology and the whole retrieval system are under development as part of a Semantic Web standardization project for the Norwegian oil and gas industry.

1 Introduction

Lots of business procedures and knowledge are written in natural language and stored in huge information repositories. The procedures are meant to support and guide employees in daily operations. The sizes of these information repositories are constantly increasing and confronting companies with a problem of efficient and effective information retrieval and reuse [7]. Retrieval of precise information at the right time is crucial in knowledge intensive tasks (e.g., monitoring performance of subsea equipment in oil and gas production). However, finding both relevant and good quality information in this sea of information is not a trivial task.

There are different strategies and aspects using IR systems. For a solution to be successful, the right problem should be tackled. In an explorative search, the user does not expect to find the information that he or she seeks at the first attempt and will probably do many searches before satisfied with the result. In these cases, the users are often uncertain about how to phrase the query and therefore will need to do a lot of reading and refining of the query before the goal is reached. Another important aspect is that most users tend to use very few terms (3 or less) in their search queries [9, 11]. Consequently, the search engine cannot *understand* the context of the user based on this little information, which result in lower precision. For instance, the term phrase “christmas tree” can be used in many different contexts, like in Christmas holiday, a component used in the oil and gas industry, etc. Furthermore, typically only the top-ranked documents in the result set are considered [9]. However, for enterprise search solutions, the users are more patient and specify their needs in more detail [9]. In particular, [9] reports that there are just about 10% of the users that are using the advanced features of a search engine, which also could have helped to improve the quality of the query. Nevertheless, at the end, the user often finds what he or she seeks but the process can be very time-consuming and frustrating.

For performance monitoring (a scenario is described in section 2), one does not have the same luxury as for explorative search. There is no time available for reading a lot of information to filter out what is relevant or not and refining the query several times. Industry needs a system that can provide correct information at the right time. Preferably, a trusted system that retrieves only highly relevant documents and where the most relevant documents are top-ranked.

An increasing number of recent information retrieval systems make use of ontologies (more details in section 6) to help the users clarify their information needs and come up with semantic representations of documents. A particular concern with these semantic approaches is integration with traditional commercial search technologies. Whereas, in this paper, we discuss how utilizing ontologies in the query process can enhance typical IR systems. In particular, we use text-mining techniques to tailor the ontology concepts to domain terminology, i.e. terms used in documents, but not necessarily aligned to standard terminology. Later, these tailored concepts are used to enrich the query to improve the retrieval quality.

Within the oil and gas industry, many companies usually have their own terminology for all the equipment available. This makes it difficult to exchange information in an efficient manner between business partners. The industry has developed a variety of ad hoc and international standards to meet this problem, though these standards mainly focus on the exchange of data between proprietary applications within a single discipline. The Integrated Information Platform for reservoir and subsea production systems (IIP) project supported by the Norwegian Research Council (NFR)[1] funds this work. The IIP project is creating an ontology for all subsea equipment used by oil and gas industry. Unlike other initiatives, this project endeavors to integrate life-cycle data spanning several standards and disciplines. A goal of this project is to define an unambiguous terminology of the domain and build an ontology that will ease integration of systems between disciplines. A common terminology is assumed to reduce risks and improve the decision making process in the industry. The project will also make this ontology publicly available and standardized by the International Organization for Standardization (ISO)[2].

The paper is organized as follows. In section 2, we present a scenario describing a typical situation where demand for accurate information requires precise query formulation. In section 3, we elucidate the settings and exemplify the ontology created in the IIP project. In section 4 and 5, we present the overall approach where we describe the architecture and techniques used. In section 6, we discuss related work. Finally, section 7 concludes the paper and discusses future work.

---

[1] http://www.forskningsradet.no

[2] http://www.iso.org/

2 Illustrative Scenario

Consider a production engineer monitoring the production efficiency of a well in the area of oil and gas exploration and production. She is located in a control room with several monitors showing the status of the wells. In such a control room, there are constant alarms of some sort with varying degree of importance. One of the most important responsibilities of the production engineer is to look for tendencies among these alarms. One or more of these alarms can indicate an upcoming serious problem that might be handled in advance and hence avoiding a potentially bigger disaster. If she can lower the risk of these potential problems by acting quickly to those relevant alarms, the production can continue smoothly and hence the company would save a lot of money. Therefore, retrieval (or delivery) of the right information at the right time is an essential task here.

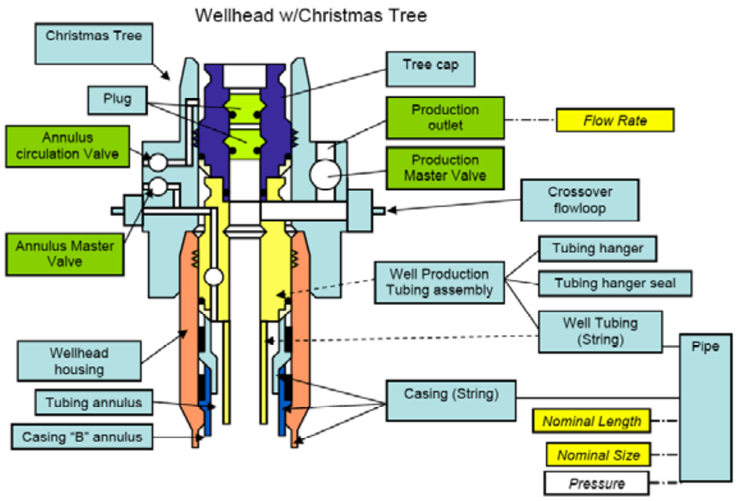

Fig. 1. A drawing of a simplified wellhead with a 'christmas tree'. Shown in the boxes are names of some of the objects and parts of a typical 'christmas tree'. Properties are in italic. [22]

Continuing the scenario, consider the production engineer noticing a tendency of some alarms that might indicate an upcoming problem. Therefore, she has to figure out if anyone has had any similar situations before and a solution to this problem. On one of her many screens she sees that one of the alarms is related to the 'production outlet', which is a part of a 'christmas tree'[3] component found among subsea equipment (see Fig. 1 for a graphical representation of a 'christmas tree'). She selects the 'production outlet' component by pointing at it on the screen. She immediately gets some status information indicating that the temperature and the pressure of the outlet is too high. The system did not find an equal situation when searching in the local document base. However, because of the ontology used in this search,

[3] "An artifact that is an assembly of pipes and piping parts, with valves and associated control equipment that is connected to the top of a wellhead and is intended for control of fluid from a well." [22]

describing the relationship of the equipment, information being relevant to this case was found after all. The system found that the 'annulus circulation valve' has a direct influence on the pressure of the 'production outlet', and therefore presents some information describing this relation. With this information presented, she adjusts the 'annulus circulation valve', and in this case prevents the potential upcoming problem.

This case illustrates the importance of information relevance and timing in the oil and gas industry. With costly and complex installations and high production volumes, the industry must make sure that the production systems are up and running at all time. Any problem that shows up needs to be solved quickly and efficiently avoiding decommissioning or waiting for the symptoms to be escalated. The engineer's task is actually even more complicated since the analysis of a particular problem may involve hundreds of potential causes and require the consultation of a large number of documents.

3 The Ontology of Reservoir and Subsea Equipment

There are many ad-hoc and international standards in the oil and gas industry. One of the main challenges for the IIP project is to integrate these standards in a unified semantic framework [14]. Fig. 2 shows the broad scope of industrial data and technical standards that are proposed for integration.

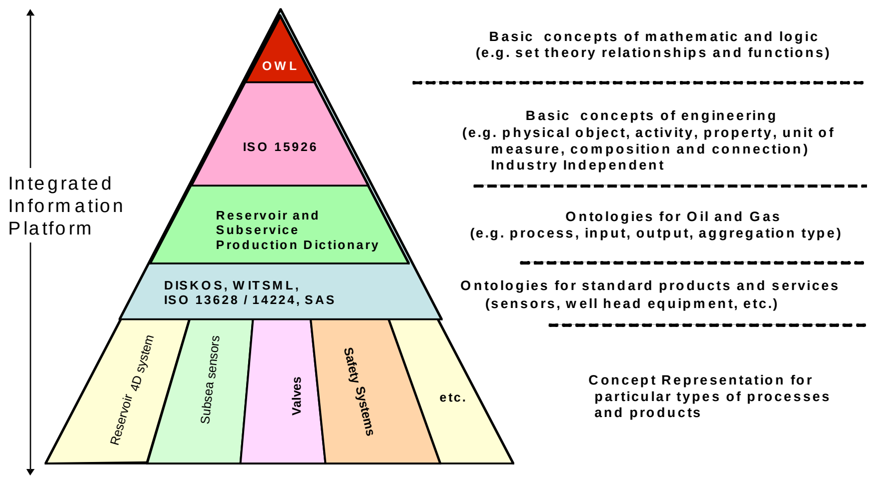

Fig. 2. The different standards being integrated into the IIP ontology [14]

The ontology created in this project is based on ISO 13628 - 'Design and operation of subsea production systems' and will be modeled in ISO 15926 – 'Integration of life-cycle data for process plants including oil and gas production facilities' Part 4 - Reference Data Library (RDL). In addition, it will include input from Statoil's Tyrihans[4] specifications, which also will serve as the test case for this project.

[4] Tyrihans is an oil and gas field in the Norwegian Sea.

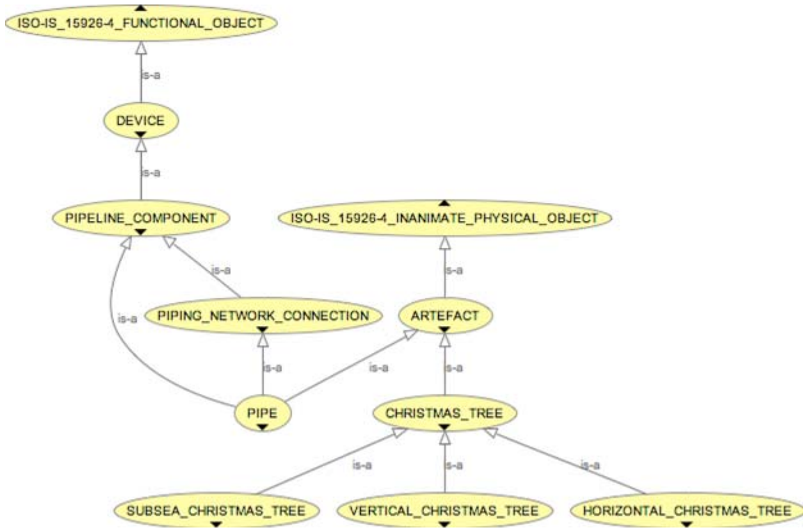

Fig. 3. Showing a fragment of the subsea equipment ontology containing ‘christmas-tree’. A graphical representation by Protégé.

Fig. 3 illustrates a fragment of the ontology under development, while Fig. 4 illustrates the concept ‘christmas tree’ represented in OWL[5]. The current state of the ontology mainly includes a hierarchical class structure, as shown in Fig. 3, and contains over 1000 subsea equipment related terms but will be extended with more terms and relations as the project proceeds.

```
<owl:Class rdf:about="#CHRISTMAS_TREE">
   …
   <dc:description
     rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
       An artefact that is an assembly of pipes and piping
       parts, with valves and associated control equipment that
       is connected to the top of a wellhead and is intended
       for control of fluid from a well.
   </dc:description>
   <dc:title
     rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
       CHRISTMAS TREE
   </dc:title>
   …
   <rdfs:subClassOf rdf:resource="#ARTEFACT"/>
</owl:Class>
```

Fig. 4. An example showing some of the information of the ‘christmas tree’ class being represented in OWL.

[5] Web Ontology Language (OWL), W3C, http://www.w3.org/2004/OWL/

4 Overall Approach

In our approach, we aim for enhanced information retrieval quality by utilizing ontologies to enrich the queries. We use the term *enrichment* in our approach, similarly as in [12], since we aim to enrich the queries to improve retrieval quality.

Fig. 5 illustrates the overall architecture of the ontology-based information retrieval system. First, we describe the individual components of the system and then the system is exemplified in the next section.

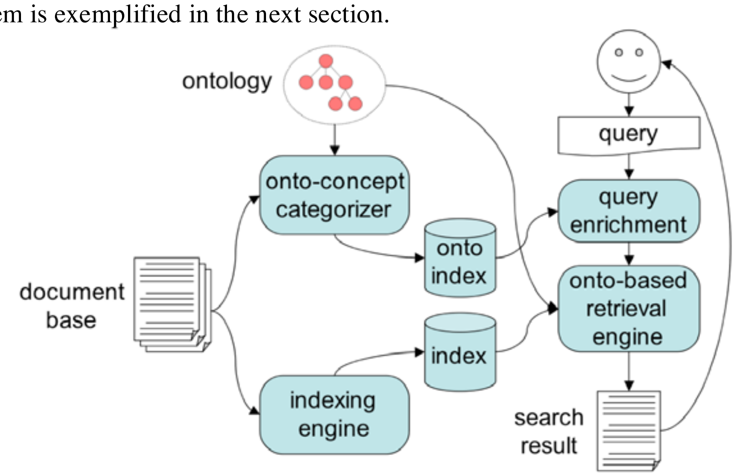

Fig. 5. The overall architecture of the ontology-based information retrieval system. The none-transparent objects illustrate the components of the system.

Onto-concept categorizer: This component extracts the terms from the document collection and associates them with relevant concept(s) from the ontology. An ontology concept is a class defined in the ontology being used. These concepts are extended into *feature vectors* with a set of relevant terms extracted from the document collection using text-mining techniques. We have adopted the k-nearest neighbor (kNN) algorithm [1, 10] to find those terms. The *feature vectors* provide interpretations of concepts with respect to the document collection and needs to be updated as the document collection changes. This allows us to relate the concepts defined in the ontology to the terms actually used in the document collection.

Indexing engine: The main task of this component is to index the document collection. The indexing system is built on top of Lucene[6], which is a freely available and fully featured text search engine from Apache[7]. Lucene is using the traditional vector space approach, counting term frequencies, and using *tf.idf* scores to calculate term weights in the index [3].

Query enrichment: This component handles the query specified by the user. The query can initially consist of concepts and/or ordinary terms (keywords). Each concept or term can be individually weighted. The concepts are replaced by corresponding feature vectors.

[6] http://lucene.apache.org/
[7] http://www.apache.org/

Onto-based retrieval engine: This component performs the search and post-processing of the retrieved results. The ontology is used when post-processing the results before presentation.

5 The Semantic Query Refinement Process

Recall the scenario in section 2. This is a typical example where concept based retrieval could be beneficial for retrieving of documents and ranking them in accordance with user's preferences. This retrieval and ranking process is important, though the result crucially hinges on the user's ability to specify unambiguously his information needs. Query interpretation is the first phase of an information retrieval session and the only part of the session that receives clear inputs from the user. Traditional vector space retrieval systems view the queries from a syntactic perspective and calculate document similarities from counting frequencies of meaningless strings. Whereas we would like the queries' *real intentions* to be exposed and reflected in the way the underlying retrieval machinery deals with them.

This system includes a query enrichment approach that uses contextually enriched ontologies to bring the queries closer to the user's preferences and the characteristics of the document collection. The idea is to associate every concept (classes and instances) of the ontology with a feature vector to tailor these concepts to the specific document collection and terminology used. Synonyms and conjugations would naturally go into such a vector, but we would also include related terms that tend to be used in connection with the concept and to provide a contextual definition of it. The weight for each term indicates to what extent the term is semantically related to the concept. For example, parts of the feature vectors of concepts 'christmas tree' and 'annulus circulation valve' are given as shown in Table 1.

Table 1. An illustration of feature vectors for two ontology concepts.

Concept	**Term that defines concept**	**Term's significance to concept**
christmas tree	christmas tree	1.0
	x-mas tree	1.0
	annulus circulation valve	0.7
	…	…
annulus circulation valve	annulus circulation valve	1.0
	christmas tree	0.5
	production outlet	0.3
	…	…

As seen from the Table 1, 'x-mas tree' is used synonymously with 'christmas tree' (weight 1.0). 'annulus circulation valve' is also closely related to 'christmas tree', though the weight of 0.7 indicates that they are not semantically as close as 'x-mas tree'. Table 1 does also show that the term 'christmas tree' has less relevance to the concept 'annulus circulation valve' (0.5) than the term 'annulus circulation valve' has to the concept 'christmas tree' (0.7). It is also worth noting that terms identical to concept names may be used to provide contextual definitions of other concepts, like in the case with 'christmas tree' and 'annulus circulation valve'.

When the user posts a query, the query is enriched by substituting the concepts in the query with corresponding feature vectors. In this way, the concepts are adopted to the terminology of the document space. For example, when the user is searching for 'christmas trees' within the oil domain he/she does not expect to get results concerning 'christmas trees' of any other domain. When all substitutions are done and new weights are calculated, the query is sent to a traditional vector space information retrieval system. This enriched query is more complex than the original one, but more precisely reflects the terminological context of concepts in the domain.

Let us now look at the way this is done formally. The query is represented as a vector Q:

$$Q = (q_1,\ldots,q_n) \tag{1}$$

where $1 \leq i \leq n$ and n is the total number of index terms in the system. Further, q_i is the normalized weight of query term i. Every concept j of the ontology is associated with a feature vector C_j so that

$$C_j = (c_{1,j},\ldots,c_{n,j}) \tag{2}$$

where c_{ij} is the normalized weight of term i for concept j. When the query Q is posted the refinement shown in Fig. 6 takes place. Q' is the semantically enriched query in terms of the user's understanding of the domain.

For all $q_i \neq 0$ in Q:
    If $C_i \neq \emptyset$:
        remove q_i from Q
        $Q' = Q + \alpha q_i C_i$

Q	*Vector for query*
Q'	*Vector of semantically enriched query*
q_i	*Weight of term i in query*
C_i	*Feature vector for concept i*
α	*Concept weight factor (default 0.6)*

Fig. 6. The query enrichment algorithm.

Consider now that the query `"CHRISTMAS TREE" "production outlet"` is posted by the user. This query is turned into the query vector $Q = (1.0_{\text{CHRISTMAS TREE}}, 1.0_{\text{production outlet}})$ before the query refinement. Whereas 'CHRISTMAS TREE' is defined to be a concept since it is capitalized, while 'production outlet' is an ordinary term. In this case, 'production outlet' also happens to be a term of 'annulus circulation valve', but that does not influence on the query refinement. By capitalizing a term, the user may decide if a term should be handled as a concept or as an ordinary term. The user can also set the weight of each concept or term for further control.

Using the ontology in IIP with the user-defined concept feature vectors shown in Table 1, we transform this query into the following enriched query:

$$Q' = (0.67_{\text{christmas tree}}, 0.67_{\text{x-mas tree}}, 0.47_{\text{annulus circulation valve}}, 1.0_{\text{production outlet}}) \quad (3)$$

This enriched query contains the original two query terms, where 'christmas tree' was defined to be a concept. The query was expanded by replacing the concept with other terms that are semantically close to the given concept of the domain. 'production outlet' ends up with a higher weight than 'christmas tree' because this is the main term of interest while the terms of the concept 'christmas tree' is given some lower weights. The effect of this is that documents being relevant to the 'christmas tree' concept that also include the term 'production outlet' will be given a higher rank.

The example shown above makes use of very small feature vectors. In practice, these feature vectors may be rather extensive, as they are automatically produced from large sets of documents. For the subsequent retrieval process, though, the contribution of terms with very low weights is negligible. Consequently, we can cut off query terms that fall below a certain threshold.

6 Related Work

Traditional information retrieval techniques (i.e., vector-space model) have an advantage of being fast and give a fair result. However, it is difficult to represent the content of the documents meaningfully using these techniques. That is, after the documents are indexed, they become a "bag of terms" and hence the semantics is partly lost in this process.

In order to increase quality of IR much effort has been put into annotating documents with semantic information [8, 12, 13, 23]. That is a tedious and labor-intensive task. Furthermore, hardly any search engines are using metadata when indexing the documents. AltaVista[8] is one of the last major search engines which dropped its support in 2002 [16]. The main reason for this is that the meta information can be and has been misused by the content providers in the purpose of giving the documents a misleading higher ranking than it should have had [16]. However, there is still a vision that for ontology based IR systems on Semantic Web, "it is necessary to annotate the web's content with terms defined in ontology" [17].

The related work to our approach comes from two main areas. Ontology based IR, in general, and approaches to query expansion, in particular. General approaches to ontology based IR can further be sub-divided into Knowledge Base (KB) and vector space model driven approaches. KB approaches use reasoning mechanism and ontological query languages to retrieve instances. Documents are treated either as instances or are annotated using ontology instances [5, 6, 17, 19]. These approaches focus on retrieving instances rather than documents. Some approaches are often combined with ontological filtering [4, 7, 15].

There are approaches combining both ontology based IR and vector space model. For instance, some start with semantic querying using ontology query languages and use resulting instances to retrieve relevant documents [19, 20]. [20] use weighted annotation when associating documents with ontology instances. The weights are

[8] http://www.altavista.com/

based on the frequency of occurrence of the instances in each document. [18] combines ontology usage with vector-space model by extending a non-ontological query. There, ontology is used to disambiguate queries. Simple text search is run on the concepts' labels and users are asked to choose the proper term interpretation. A similar approach is described in [25] where documents are associated with concepts in the ontology. The concepts in the query are matched to the concepts of the ontology in order to retrieve terms and then used for calculation of document similarity.

[4] is using ontologies for retrieval and filtering of domain information across multiple domains. There each ontology concept is defined as a domain feature with detailed information relevant to the domain including relationships with other features. The relationships used are hypernyms (super class), hyponyms (sub class), and synonyms. Unfortunately, there are no details in [4] provided on how a domain feature is created.

Most query enrichment approaches are not using ontologies like [2, 24, 26]. Query expansion is typically done by extending provided query terms with synonyms or hyponyms (cf. [21]). Some approaches are focusing on using ontologies in the process of enriching queries [4, 6, 25]. However, ontology in such case typically serves as thesaurus containing synonyms, hypernyms/hyponyms, and do not consider the context of each term, i.e. every term is equally weighted.

[24] is using query expansion based on similarity thesaurus. Weighting of terms is used to reflect the domain knowledge. The query expansion is done by similarity measures. Similarly, [2] describes a conceptual query expansion. There, the query concepts are created from a result set. Both approaches show an improvement compared to simple term based queries.

While an approach presented in [26] is most similar to ours. However, [26] is not using ontologies but is reliant on query concepts. Two techniques are used to create the feature vectors of the query concepts, i.e. based on document set and result set of a user query.

To contrast to above discussed related work we emphasize the main features of our approach as follows. Our approach relies on domain knowledge represented in ontology when constructing feature vectors, then traditional vector-space retrieval model is used for the information retrieval task, where feature vectors are used to enrich provided queries. The main advantage of our approach is that the concepts of an ontology is tailored to the terminology of the document collection, which can vary a lot even within the same domain.

7 Conclusions and Future Work

In this paper, we have proposed a method for utilizing ontologies to improve the retrieval quality. The concepts in the ontology are associated with contextual definitions in terms of weighted feature vectors tailoring the ontology to the content of the document collection. Further, the feature vectors are used to enrich a provided query. Query enrichment by feature vectors provides means to bridge the gap between query terms and terminology used in a document set, and still employing the knowledge encoded in the ontology. We have also proposed that concepts and ordinary terms or keywords of the query should be handled differently since they have different roles.

Main architectural components and techniques constituting the method have been presented in this paper. As research reported here is still in progress, we have not been able to formally evaluate the approach. Though, preliminary results indicate that the quality of the feature vectors is very important for the quality of the search result.

In future work we are planning to inspect and tackle a set of issues as follows. First, there is a need to refine term weight computation. Here we will investigate alternative methods for assigning relevant terms to the ontology concepts, i.e. using association rules, and evaluate the influence on the search results. Second, we will also look into alternative methods for post-processing of the results utilizing the semantic relations in the ontology for better ranking and navigation.

Acknowledgements. This research work is funded by the Integrated Information Platform for reservoir and subsea production systems (IIP) project, which is supported by the Norwegian Research Council (NFR). NFR project number 163457/S30.

References

1. Aas, K., Eikvil, L.: Text categorisation: a survey. Technical report, no. 941. Norwegian Computing Center, Oslo (1999) 37 p.
2. Grootjen, F.A., van der Weide, T.P.: Conceptual query expansion. Data & Knowledge Engineering 56 (2006) 174-193
3. Baeza-Yates, R., Ribeiro-Neto, B.: Modern information retrieval. ACM Press, New York (1999)
4. Braga, R.M.M., Werner, C.M.L., Mattoso, M.: Using Ontologies for Domain Information Retrieval. Proceedings of the 11th International Workshop on Database and Expert Systems Applications. IEEE Computer Society (2000) 836-840
5. Rocha, C., Schwabe, D., de Aragao, M.P.: A hybrid approach for searching in the semantic web. Proceeding of WWW 2004, ACM (2004) 374-383
6. Ciorăscu, C., Ciorăscu, I., Stoffel, K.: knOWLer - Ontological Support for Information Retrieval Systems. In Proceedings of Sigir 2003 Conference, Workshop on Semantic Web, Toronto, Canada (2003)
7. Borghoff, U.M., Pareschi, R.: Information Technology for Knowledge Management. Journal of Universal Computer Science 3 (1997) 835-842
8. Desmontils, E., Jacquin, C.: Indexing a Web Site with a Terminology Oriented Ontology. In I.F. Cruz, S. Decker, J. Euzenat and D.L. McGuinness (eds.) The Emerging Semantic Web. IOS Press, (2002) 181- 198
9. Gulla, J.A., Auran, P.G., Risvik, K.M.: Linguistic Techniques in Large-Scale Search Engines. Fast Search & Transfer (2002) 15 p.
10. Mitchell, T.M.: Machine learning. McGraw-Hill, New York (1997)
11. Spink, A., Wolfram, D., Jansen, M.B.J., Saracevic, T.: Searching the Web: the public and their queries. J. Am. Soc. Inf. Sci. Technol. 52 (2001) 226-234
12. Motta, E., Shum, S.B., Domingue, J.: Case Studies in Ontology-Driven Document Enrichment: Principles, Tools and Applications. International Journal of Human-Computer Studies 6 (2000) 1071-1109
13. Popov, B., Kiryakov, A., Kirilov, A., Manov, D., Ognyanoff, D., Goranov, M.: KIM - Semantic Annotation Platform. In: Fensel, D., Sycara, K.P., Mylopoulos, J. (eds.): The

Semantic Web - ISWC 2003, Second International Semantic Web Conference, Sanibel Island, FL, USA, October 20-23, 2003, Proceedings, Vol. 2870. Springer (2003) 834-849

14. Sandsmark, N., Mehta, S.: Integrated Information Platform for Reservoir and Subsea Production Systems. (2004) 9 p.
15. Shah, U., Finin, T., Joshi, A., Cost, R.S., Mayfield, J.: Information Retrieval On The Semantic Web. Proceedings of Conference on Information and Knowledge Management. ACM Press, McLean, Virginia, USA (2002) 461-468
16. Sullivan, D.: Death of a Meta Tag. Search Engine Watch (2002)
17. Song, J-F., Zhang, W-M., Xiao, W., Li, G-H., Xu, Z-N.: Ontology-Based Information Retrieval Model for the Semantic Web. Proceedings of EEE 2005. IEEE Computer Society (2005) 152-155
18. Nagypal, G.: Improving Information Retrieval Effectiveness by Using Domain Knowledge Stored in Ontologies. OTM Workshops 2005, LNCS 3762, Springer-Verlag, (2005) 780-789
19. Kiryakov, A., Popov, B, Terziev, I., Manov, D., and Ognyanoff, D.: Semantic Annotation, Indexing, and Retrieval. Journal of Web Semantics 2(1), Elsevier, (2005)
20. Vallet, D, Fernández, M., Castells, P.: An Ontology-Based Information Retrieval Model. Gómez-Pérez, A., Euzenat, J. (Eds.): Proceedings of ESWC 2005, LNCS 3532, Springer-Verlag. (2005) 455-470.
21. Chenggang, W., Wenpin, J., Qijia, T. et al.: An information retrieval server based on ontology and multiagent. Journal of computer research & development 38(6) (2001) 641-647.
22. DNV: Tyrihans Terminology for Subsea Equipment and Subsea Production Data. DNV (2005) 60 p.
23. Fensel, D., Harmelen, F.v., Klein, M., Akkermans, H., Broekstra, J., Fluit, C., Meer, J.v.d., Schnurr, H.-P., Studer, R., Hughes, J., Krohn, U., Davies, J., Engels, R., Bremdal, B., Ygge, F., Lau, T., Novotny, B., Reimer, U., Horrocks, I.: On-To-Knowledge: Ontology-based Tools for Knowledge Management. In Proceedings of the eBusiness and eWork 2000 (EMMSEC 2000) Conference, Madrid, Spain (2000)
24. Qiu, Y., Frei, H.-P.: Concept based query expansion. Proceedings of the 16th annual international ACM SIGIR conference on Research and development in information retrieval. ACM Press, Pittsburgh, Pennsylvania, USA (1993) 160-169
25. Paralic, J., Kostial, I.: Ontology-based Information Retrieval. Information and Intelligent Systems, Croatia (2003) 23-28
26. Chang, Y., Ounis, I., Kim, M.: Query reformulation using automatically generated query concepts from a document space. Information Processing and Management 42 (2006) 453-468