

KNOWLEDGE REPRESENTATION FOR SUPPLIER DISCOVERY IN DISTRIBUTED MANUFACTURING

Christian McArthus, Farhad Ameri
Department of Engineering Technology
Texas State University
San Marcos, TX 78666

ABSTRACT

Online outsourcing has recently gained popularity among small and medium sized manufacturing companies as an efficient method for building flexible network of manufacturing counterparts. Several electronic marketplaces have emerged within the last few years with the objective of enabling large communities of buyers and sellers to virtually meet and establish new partnerships. Although e-marketplaces typically provide different automated search capabilities, they mainly rely on human users for final screening and evaluation of qualified suppliers. As the size of supply and demand pools increase, human-based search becomes inefficient. This paper describes an effort for enhancing the automation capabilities of web-based markets through an ontological approach. The proposed ontology is referred to as Manufacturing Service Description Language. MSDL provides formal semantic for manufacturing knowledge representation, thus enabling machine agents to actively participate in supplier discovery process.

Keywords: Manufacturing outsourcing, ontology, semantic search

1. INTRODUCTION

Web-based marketplaces for manufacturing services are currently the state-of-the-art in developing flexible supply networks for discrete part manufacturing [1, 2]. The growing popularity of online markets for manufacturing services can be attributed to several factors such as low cost of entrance, low cost of transaction due to elimination of market mediators, the possibility of interacting with a far larger number of potential counterparts, and equal treatment of members regardless of their size and global reach[3]. ChemConnect¹ and Covisint² are examples of online markets, in the chemical and automotive industries respectively, which evolved rapidly within a few years and now provide e-commerce services for several thousands of companies in more than 150 countries.

The general architecture of online markets for manufacturing services is depicted in Figure 1. Using the bidding system, sellers bid on selected RFQs (request for quote) available in the RFQ database. Buyers evaluate the bids and select the ones which best meet their requirements. Online RFQ markets usually provide traditional means of search including keyword search, directory search, and database search on both buyer and seller data. Search criteria for querying the buyers database typically include customer's category (process, sub-process, and product), materials, delivery location, industry, quantities (min-max), and dimensions. On the seller's side, search criteria often include process, sub-process, company name and location, and quality certification. These criteria, however, are simplistic and often provide incomplete picture of supplier's potential match with requirements, leading to identification of suppliers that are irrelevant. Therefore, to arrive at more accurate results, the output of the search engine is further refined by human users through reviewing the narrative description of suppliers' capabilities provided in a free-text format. However, as the size of the search space increases, human-based evaluation and screening becomes increasingly inefficient and error-prone.

¹ <http://www.chemconnect.com/>

² <http://www.covisint.com/>

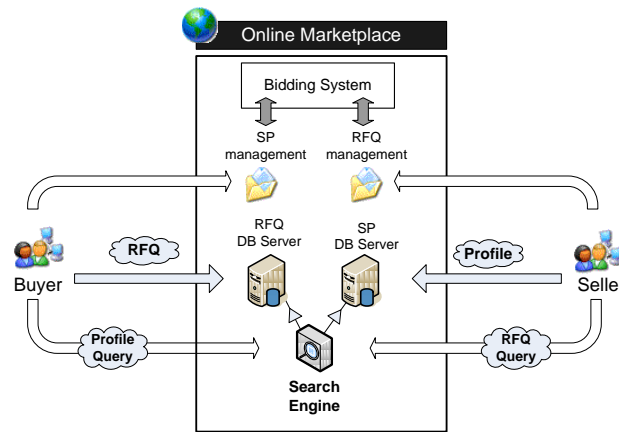


Figure 1. General architecture of online RFQ markets

To enable automated search and evaluation with high precision, search engines in RFQ markets need to be equipped with semantic search capabilities. The search tools that operate in the existing markets rely on the syntax of the content and not the semantics. In a syntactic search, the results very much depend on the choice of keywords used in the query. Therefore, relevant documents that use different syntax may be excluded from the search results (false negative error) and similarly, irrelevant documents with similar syntax with the query may be deemed relevant (false positive error) by the search engine.

Semantic search capability highly depends on the underlying information modeling formalism used in the market. Object-oriented models, that are the most widely used models in online market, represent a domain of interest merely through objects and properties of those objects. There is no mechanism for defining logical constraints that can further represent the semantics of each object. Therefore, object-oriented models lack enough expressivity for describing manufacturing domain knowledge in a rich and machine-interpretable fashion. Additionally, due the rigid structure of the object-oriented models, expansion of the knowledge base is not readily possible.

To address the identified gaps in the underlying information model of online RFQ markets, an ontological approach is adopted in the current research. A formal ontology provides explicit representation of information semantics, thus enabling autonomous reasoning and inference by machine agents. In presence of a rich information model, more intelligent search algorithms can be employed to improve the precision of the returned results. Researchers have used ontological approach for dealing with semantic interoperability in distributed manufacturing [4-7]. However, the existing ontologies do not have sufficient expressivity for addressing the supplier discovery problem.

The proposed ontology in this work is coupled with a set of intelligent search algorithms that can quantify the semantic proximity of supply and demand entities. The search algorithms are designed such that they can return partial matches based on the semantic similarities of the advertisement and the query. The open-source architecture of the ontology will allow for customization of the search logics based on particular scenarios. The extensible nature of the ontology will allow suppliers to freely describe their capabilities using a wider range of concepts without being confined to a limited set of terms provided by the RFQ market. Similarly, customers will have flexibility for formulating their queries and can emphasize on a certain aspect of their desired service such as specialized machine tools, quality control services, or packaging systems. One of the interesting aspects of a shared ontology is that ontology users can extend the ontology as they use it. In this way, the ontology will expand and evolve over time to include new terms and rules that form the manufacturing domain knowledge.

The remainder of this paper is organized as follows. The next section provides an overview of the proposed ontology. The user interfaces required for demand description are discussed afterwards followed by a description of a feature recognition system used for identification of manufacturing

requirements of parts based on their CAD model. The paper ends with a brief description of the similarity measurement techniques used in this work together with a discussion on the results of an experimental evaluation of the proposed framework.

2. MANUFACTURING SERVICE DESCRIPTION LANGUAGE (MSDL)

As mentioned earlier, to improve the intelligence of automated supplier discovery systems, the expressivity of their underlying information model should be enhanced. One promising solution is adaptation of an ontological approach for supply and demand representation. To this end, an ontology is proposed in this work with the objective of semantic description of *manufacturing services*. The proposed ontology is referred to as the Manufacturing Service Description Language (MSDL). OWL-DL, a sub-language of OWL, is selected as the ontology language of MSDL. OWL is recommended by World Wide Web Consortium (W3C) as the ontology language of the Semantic Web. OWL uses XML as the syntax language; hence it has enough portability, flexibility, and extensibility for web-scale applications. Description Logic (DL)[8] is supported by the Semantic Web (SW) meaning that OWL-based ontologies can be shared, parsed, and manipulated through open-source web-based tools and technologies. The underlying knowledge modeling formalism of OWL is DL. DL provides formal syntax and semantics for developing information and knowledge models with a domain of interest in terms of concepts, relationships between concepts and logical constraints that concepts must satisfy. Since MSDL's semantics are mainly in the form of concept definitions and their interrelations (taxonomies), the expressivity offered through DL is sufficient for knowledge representation in this work. Due to its mathematically rigorous formalism, DL supports automated reasoning services such as concept subsumption, concept equivalence, and concept consistency. The scope of MSDL, in the initial development is limited to conventional *machining services* such as turning, drilling, and milling.

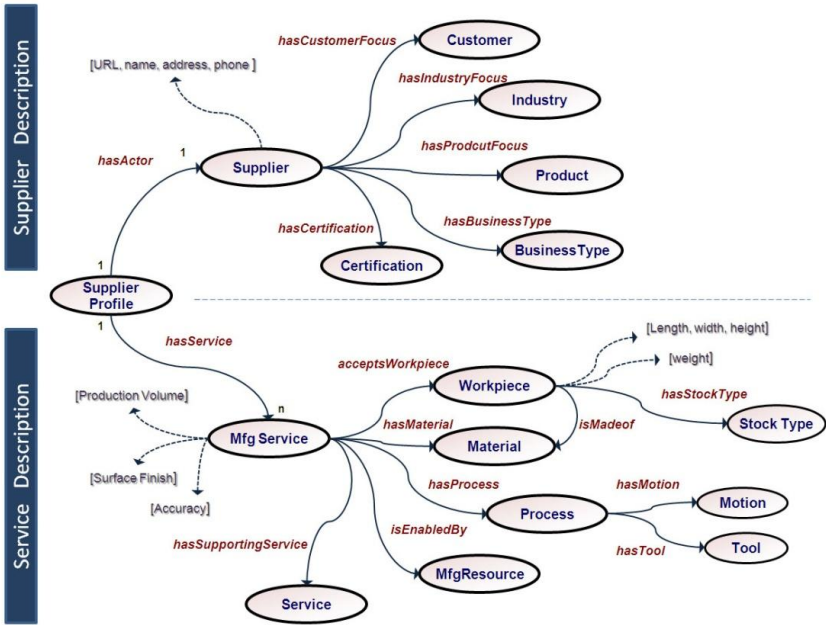


Figure 2. Concept diagram for Supplier Profile class

In MSDL, supply and demand are represented by Supplier Profile and RFQ classes respectively. As can be seen in Figure 2, Supplier Profile has two major components, namely, the supplier and the manufacturing services that the supplier provides. Services are further described through their associated processes, materials, resources, and supporting services. Figure 3 shows the concept diagram for RFQ. As shown in this figure, each RFQ has exactly one Customer. Also, an RFQ can have multiple services. Each RFQ has a *Part* class connected to it through the *hasPart* property. The *Part* class is used for defining the attributes of the part contained in the query. The *Customer* and *Supplier* classes are both sub-classes of the *Actor* class, an imported class from the OWL-S³ ontology.

³ <http://www.ai.sri.com/daml/services/owl-s/1.0/>

The *Actor* class contains the generic attributes required for introducing a supplier including name, physical address, phone, fax, and web URL. However, more specific attributes such as industry focus, product focus, company size, quality certifications, and years of experience, which collectively define the actor's profile, are directly connected to the Supplier and Customer classes.

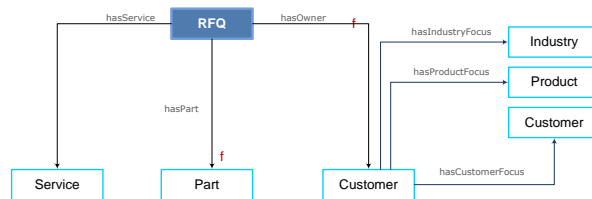


Figure 3. Concept Diagram for RFQ in MSDL

There are two major methods for encoding further semantics (beyond concepts and properties) in MSDL. The first method is building taxonomies (i.e., explicit parent-child relationships) and the second method is using formal definition of classes. For example, the semantics of Industry and Product types are encoded in the form of explicit taxonomies based on North American Industry Classification System⁴ (NAICS) and United Nation Central Product Classification⁵ (UN CPC). Concepts such as *Process* and *Material* are formally defined through necessary and sufficient conditions. For example, Figure 4 provides the formal definition of the *end milling* process through the major feed and cutting motions involved in this process. Since concepts such as Motion, Tool, Machining, and Axis have their own formal definition software agents can understand and interpret the *meaning* of the end milling process as opposed to merely relying on its name.

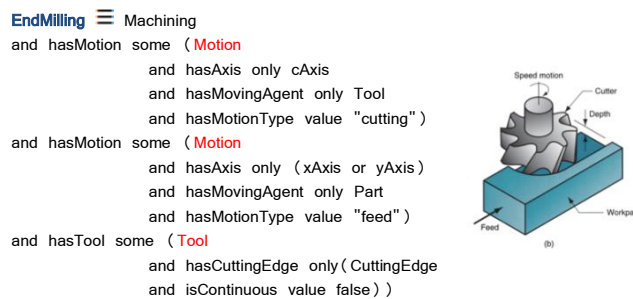


Figure 4. Formal definition of end milling process (picture courtesy of John Wiley & Sons Inc.)

3. RFQ GENERATION

In order for the users to describe their capabilities and needs in terms of manufacturing services using MSDL representation, two types of interfaces are required, one for customers and one for suppliers. This paper focuses on the first type of interfaces that aid customers in accurately defining their manufacturing requirements via RFQs. The generated RFQs can be maintained in a central directory and suppliers' search agents can query the directory and find the appropriate matches. Since ontology editors such as Protégé⁶ are designed at the outset to be used by the ontology developers who are familiar with the technical aspects of the DL ontologies, it is necessary to provide the users with more user-friendly and intuitive interfaces. Figure 5, Figure 6, and Figure 7 show examples of a web-based interface developed for creating RFQ instances. The developed interfaces hide the complexities of the underlying knowledge-based model and provide the users with intuitive, yet comprehensive templates for describing their capabilities and needs.

⁴ <http://www.census.gov/eos/www/naics/>

⁵ <http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=16>

⁶ <http://protege.stanford.edu/>

Customer Profile Information

Company Name Website URL
 Phone Number Fax Number
 Rating
 Business Type
 Industry Focus Product Focus
 Choose Industries Choose Products
 Customer Focus
 Production Range Certification

Figure 5. The GUI for customer information

Figure 5 shows the interface for collecting customer information. Apart from general information such as name, business type, phone number, and web address, customers can provide information on their particular industry and type of products they produce. Industry and product focus are two key features of similarity based on the proposed search algorithms. After providing general information about the customer, the next step is to define the part and the manufacturing services that the part requires. Since a considerable amount of technical information about a part is encapsulated in the CAD file of the part, it is a reasonable approach to extract the part-related information automatically from the CAD file. Therefore, as can be seen in Figure 6, customers can upload a part's CAD file when they are providing information about the part corresponding to a specific RFQ. In particular, from the CAD file, it is possible to identify the required machining processes. To this end, a Feature Recognition (FR) module is incorporated in the proposed system that identifies manufacturing features of a given part and maps them to relevant manufacturing processes. For example, if the CAD model contains at least one *hole feature* it can be inferred that the part requires a *drilling* service. Dimensional information such as the bounding box of the part as well as dimensions of each feature can also be obtained from the CAD file and incorporated in formulation of the RFQ. Dimensional information is necessary to verify if the suppliers' machine tools can accommodate the part and provide the required tolerances.

The next interface, shown in Figure 7, deals with service information. If there is no CAD file available for a part, customers can directly select the necessary manufacturing processes from the taxonomy of processes available in MSDL. Also, through this interface, it is possible to identify the desirable supporting services such packaging, heat treatment, or quality control. Furthermore, customer can describe the specific machine tools they require for the service, the type of material the service should be able to support, as well as the type of workpiece the service should accept.

3.1. Automated Recognition of Manufacturing Processes

The integration of computer-aided design (CAD) and computer-aided manufacturing (CAM) systems has been one of the major areas of focus for researchers involved in manufacturing domain for decades. Automatic feature recognition techniques were originally developed to bridge the gap between CAD and CAM systems through translating geometric data into higher level information with manufacturing significance. This paper reports one of the first applications of automatic feature recognition in manufacturing supplier selection. In particular, this work proposes a platform-dependant technique for identifying the manufacturing requirements of mechanical part based on its CAD model. In the platform-independent approach, the part's geometrical data are extracted from a neutral file such as DXF, IGES, or STEP. In contrast, the platform-dependent approach extracts the information of the design features directly from a design-by-feature solid model through the object-oriented model of a part. In the former approach which uses a form of intermediate CAD file format

(e.g., DXF, STL, STEP, etc.), the data processing time and error are increased. However, the platform dependant approach eliminates these deficiencies through eliminating the intermediate data files. In the platform-dependent approach, both the creation of the CAD part model and feature recognition are done in the same CAD platform. Some of the advantages of this approach are:

- Easy access to the vast objects library of design features.
- Ability to obtain and organize the geometry and topology entities of parts as an object-oriented model.
- Exploitation of the built-in functions (API) from within the solid modeler in order to facilitate the application of the feature recognition rules.



Request For Quote System

Part Information

You may upload a CAD file for the part you are creating the RFQ for.

Diameter

Length

Width

Product Type

Shape

Material

Figure 6. The GUI for general RFQ information



Request For Quote System

Service Information

Process

Supporting Services

Enabled By

Materials

Workpiece(s)

Production Volume

Figure 7. The GUI for service information

The ability to handle non-intersecting and intersecting features recognition based on design-by-feature approach was the focus of recent research efforts. In such an approach, direct recognition of

manufacturing features from a CAD or a solid model is identified as a necessary component of an automated design analysis [9-11]. Autodesk Inventor is selected as the platform for the feature recognition system developed for supplier selection purpose. Each recognized feature provides critical data (e.g. length, width, depth, vertexes, tools accessing direction, the volume of feature, etc) that can be used in evaluating the manufacturing requirements of a given part. Pocket, step, slot, hole, counterbore, and countersink are among the features that will be identified by the developed program.

A designer who is interacting with a solid modeler may choose two different methods for creating a design feature such as hole/Counterbore/Countersink. The first method is the use of predefined built-in objects whose behaviors are identical to their feature kinds. The second method is to create a 2D profile including a circle entity and then sweeping it in the negative direction. The designed software in this study is capable of functioning with models made by either of the two methods. The critical information (e.g. vertices, volume, the type of surfaces, tools accessing direction) of the recognized features were also extracted and calculated while identifying those features. The proposed algorithm for the feature recognition is elaborated in Figure 8. After the extraction of manufacturing information; they are automatically transferred from Autodesk Inventor to MS Access. The Access tables are then reformatted into structured text files and parsed programmatically to translate their contents into MSDL format. Figure 9 shows a sample part together with the extracted features and dimensional information. In this example, since the recognized features are all prismatic, they are mapped to milling and EDM processes.

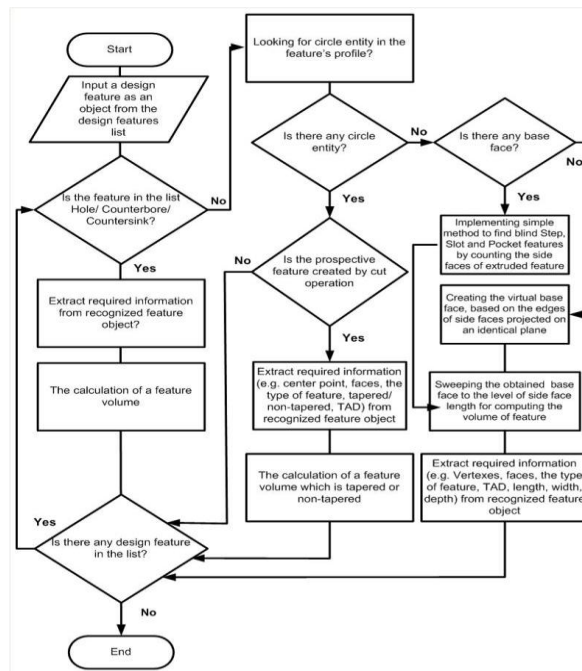


Figure 8. Flowchart for feature recognition in platform-dependent area

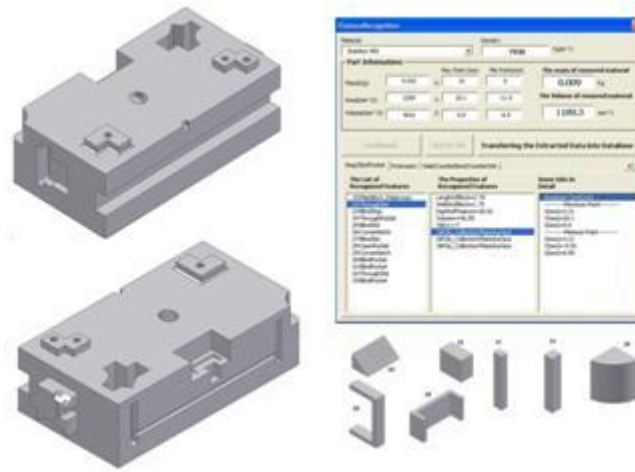


Figure 9. Illustration of sample part with its recognized features (slot, step, pocket, corner, and corner notch) with their manufacturing data

4. SIMILARITY MEASUREMENT ALGORITHMS

The search space of the RFQ market consists of the profiles of the suppliers who provide various types of manufacturing service together with RFQs that encode demand entities. The goal is to quantify the similarities of each profile in the search space with a particular service request formulated as an RFQ. The similarity of a supplier profile (SP) to a given query (Q) is calculated through the weighted average of the similarities of the actor (i.e., the supplier who provides the service) and service portions of the query, as shown by the following equation:

$$Sim(Q, SP) = Sim_{actor} + Sim_{service}$$

The actor and service similarity scores are themselves weighted averages of their respective elements. The actor similarity score (Sim_{actor}) is a measure of the similarity of the desirable supplier described in the query and the actual supplier who owns the profile of interest. Actor similarity score is composed of three elements, namely, industry, product, and certification.

$$Sim_{actor}(Q, SP) = w_{ind}Sim_{ind} + w_{prod}Sim_{prod} + w_{cert}Sim_{cert}$$

Since MSDL contains explicit classifications of industries and products, the similarity scores for industry and product elements of the query are determined through a taxonomy-based approach as shown by the following equation[12]:

$$Sim(C_A, C_B) = \frac{2 \log(IC_{sa})}{\log(IC_A) + \log(IC_B)}$$

Where IC_A and IC_B are the information content (IC) of the two nodes being compared and IC_{sa} is the information content of their shared ancestor, or common parent, in the taxonomy tree.

The service similarity score ($Sim_{service}$) for a query and profile pair is measured along five dimensions specified in the query: processes, materials, required supporting services, production volume, and required precision.

$$Sim_{service}(Q, SP) = w_{proc}Sim_{proc} + w_{mat}Sim_{mat} + w_{SuppSer}Sim_{SuppSer} + w_{Vol}Sim_{Vol} + w_{prec}Sim_{prec}$$

In the MSDL, process and material are formally defined through a set of necessary and sufficient conditions (features). These conditions provide a semantic description of their associated concepts. Therefore, a feature-based method can be employed for measuring the similarities of instances of process or material classes through a comparison of each item's conditions:

$$Sim(A, B) = \frac{n_{A \cap B}}{n_{A \cap B} + \mu n_{A-B} + \nu n_{B-A}}$$

Where $\eta_{A \cap B}$ is the number of conditions common to both the query and the search space class, η_{A-B} is the number of conditions in the query that are not in the search space class, and η_{B-A} is the number of conditions in the search space that are not in the query class. The variables μ and ν are weighting factors to allow for stressing the importance of the query or the search space. A query can contain more than one service. Therefore, the service similarity score can be measured either based on maximum similarity between the requested services and the provided ones or it can be based on the average service similarity obtainable in a query-profile pair.

5. EXPERIMENTAL EVALUATION

To evaluate the potentials of the proposed matching technique in terms of correlation with human judgment, a series of experiments were designed and conducted based on real-life data. A sample of 30 supplier profiles was collected from Thomas Net⁷, a leading web-based portal for manufacturing supplier outsourcing to form the base collection for the experiments. Supplier profiles in Thomas Net are basically paragraphs of text describing the services the supplier offer, workable materials, a brief description of the machineries and the equipment, as well as product and industry focus of the supplier. These textual descriptions were manually converted into MSDL vocabulary in order to accurately describe the supplier capabilities semantically. Two real RFQs were selected from Mfg.com⁸, a major RFQ market in manufacturing domain. Figure 10 shows one of the selected RFQs. The selected RFQs were given to the domain experts along with the textual descriptions of the thirty suppliers in the base collection and they were asked to rank the suppliers according to their suitability for producing the part described in each RFQ. The textual descriptions were also searched using a keyword search method. Finally, the RFQs were converted into MSDL and run on the MSDL search engine. The null hypothesis for this experiment was that there is no meaningful difference between the results obtained from the keyword search methods and the results returned by the MSDL search method whereas, the alternative hypothesis claimed a meaningful difference.

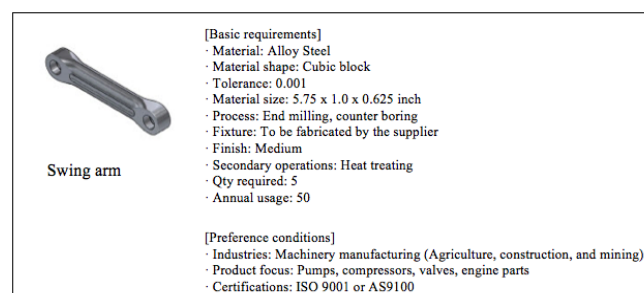


Figure 10. Swing arm RFQ information

As a measure of performance for the competing search methods, Discounted Cumulative Gain (DCG)[13] was employed. DCG quantifies the usefulness or gain of search results by examining both the relevance of the returned suppliers as well as the order by which they are returned. The rankings produced by the domain experts were considered to be the ideal results and were used as the reference. Since two experts were using different approaches for supplier selections, two independent DCG analysis were performed based each expert. **Figure 11** summarizes the final results obtained through the experimentations. In this figure, the vertical axis represents the normalized DCG value.

As can be seen in this figure, MSDL search method outperforms keyword-based search method in three out of four cases. Particularly in the second query which contains more detailed information. The results obtained through the MSDL search engine demonstrate a meaningful improvement over the keyword method. The reason is that as the complexity of the query increases, so does the level of required reasoning and inference for arriving at useful results. A study of the online RFQ markets revealed that a typical RFQ is more complex than the queries used in this experiment, which makes the manual search extremely time-consuming and inefficient especially when the size of search space

⁷ <http://www.thomasnet.com/>

⁸ <http://www.mfg.com/>

grows. Overall, the findings of this experiment were in favor the alternative hypothesis that suggests ontology-based search yields more accurate results as compared to the keyword search method.

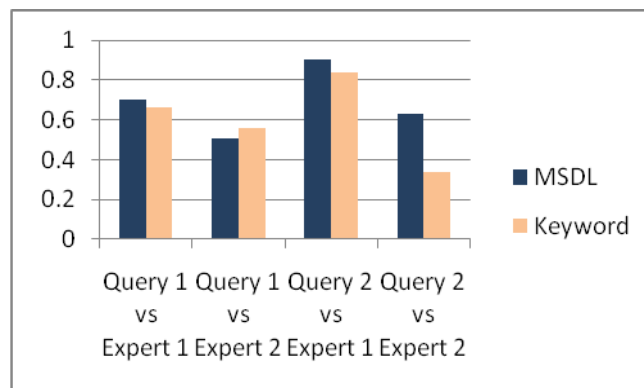


Figure 11. MSDL vs. Keyword search comparison

6. CONCLUSIONS

Timely configuration of supply chains with required capabilities in distributed environments calls for precise and rapid evaluation of potential partners. Existing web-based solutions for manufacturing partner selection highly rely on human users for screening and evaluating potential partners. To improve the intelligence of search engine in RFQ markets, an ontological approach was proposed in this work. Through incorporation of logic-based semantics, MSDL enables unambiguous description of supply and demand entities in terms of manufacturing services. Also, due to its formality, MSDL enables active involvement of machine agents in supply chain configuration process. Development of MSDL is a step toward enhancing the automation capabilities of current online markets for manufacturing services. With a knowledge-based supply chain configuration system, not only will the duration of supply chain formation be reduced considerably, but also the desirability of the resulting supply chain will be increased drastically. Through experimental evaluation of the proposed search algorithms, it was concluded that the semantic search method outperforms the keyword search method. The feature recognition module embedded in the web-based system allows for automated formulation of queries based on manufacturing requirements of parts. The web-based interfaces introduced in this work are the initial modules of a more comprehensive web-based system that is under development which will be used as a test-bed for evaluating the proposed semantic search and evaluation methodology based on actual data collected from contract manufacturers in central Texas area. The existing version of MSDL is based on machining processes. However, its underlying concepts are applicable to all manufacturing processes such as forming, casting, and joining. Currently the reasoning services used in the search engine is limited to subsumption checking. However, to reap the real benefits of a formal ontology, more advance reasoning techniques should be incorporated in the search engine. Other future works in this area include expansion of MSDL to other manufacturing and development of a multi-agent system for supply chain deployment.

REFERENCES

1. Akarte, M.M., et al., *Web baed casting supplier evaluation using analytical hierarchy process*. Journal of Operational Reaserch Society, 2001. **52**(5).
2. Emiliani, M.L., *Sourcing in the global aerospace supply chain using online reverse auctions*. Industrial Marketing Management, 2004. **33**(1): p. 65-72.
3. Das, S.K. and H. Shahin, *MODELS FOR SUPPLY CHAIN VENDOR SELECTION IN E-MARKETS*. Journal of the Chinese Institute of Industrial Engineers, 2003. **20**(3): p. 231 - 239.
4. Lin, H.K. and J.A. Harding, *A manufacturing system engineering ontology model on the semantic web for inter-enterprise collaboration*. Comput. Ind., 2007. **58**(5): p. 428-437.
5. Jang, J., et al., *Discovering and integrating distributed manufacturing services with semantic manufacturing capability profiles*. Int. J. Comput. Integr. Manuf., 2008. **21**(6): p. 631-646.
6. Kulvatunyou, B., H. Cho, and Y.J. Son, *A semantic web service framework to support intelligent distributed manufacturing*. Int. J. Know.-Based Intell. Eng. Syst., 2005. **9**(2): p. 107-127.

7. Zhonghua, Y., J.B. Zhangx, and L. Chor Ping. *Towards Dynamic Integration of Collaborative Virtual Enterprise Using Semantic Web Services*. in *Industrial Informatics, 2006 IEEE International Conference on*. 2006.
8. Baader, F., *The description logic handbook: Theory, implementation and applications* 2003: Cambridge University Press.
9. Hayasi, M. and B. Asiabanpour, *Extraction of manufacturing information from design-by-feature solid model through feature recognition*. *The International Journal of Advanced Manufacturing Technology*, 2009. **44**(11): p. 1191-1203.
10. Yeol Lee, J. and K. Kim, *A feature-based approach to extracting machining features*. *Computer-Aided Design*, 1998. **30**(13): p. 1019-1035.
11. Woo, Y. and H. Sakurai, *Recognition of maximal features by volume decomposition*. *Computer-Aided Design*, 2002. **34**(3): p. 195-207.
12. Lin, D., *An Information-Theoretic Definition of Similarity*, in *Proceedings of the Fifteenth International Conference on Machine Learning*. 1998, Morgan Kaufmann Publishers Inc.
13. Jarvelin, K. and J. Kekalainen, *IR evaluation methods for retrieving highly relevant documents*. *SIGIR Forum (ACM Special Interest Group on Information Retrieval)*, 2000: p. 41-48.