Proceedings of the ASME 2011 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference
IDETC/CIE 2011
August 28-31, 2011, Washington, DC, USA

DETC2011-48520

TOWARDS INDUSTRIAL IMPLEMENTATION OF EMERGING SEMANTIC TECHNOLOGIES

Jay T. Breindel, Ian R. Grosse, Sundar Krishnamurty, John Altidor, Jack Wileden
University of Massachusetts Amherst
Amherst, MA 01003
jbreinde@engin.umass.edu, {grosse,skrishna}@ecs.umass.edu, {jaltidor,jack}@cs.umass.edu

Seth Trachtenberg
Raytheon Integrated Defense Systems
Sudbury, MA 01776
Seth_Trachtenberg@raytheon.com

Paul Witherell
National Institute of Standards and Technology
Gaithersburg, MD 20899
paul.witherell@nist.gov

ABSTRACT

Every new design, project, or procedure within a company generates a considerable amount of new information and important knowledge. Furthermore, a tremendous amount of legacy knowledge already exists in companies in electronic and non-electronic formats, and techniques are needed for representing, structuring and reusing this knowledge. Many researchers have spent considerable time and effort developing semantic knowledge management systems, which in theory are presumed to address these problems. Despite significant research investments, little has been done to implement these systems within an industrial setting.

In this paper we identify five main requirements to the development of an industry-ready application of semantic knowledge management systems and discuss how each of these can be addressed. These requirements include the ease of new knowledge management software adoption, the incorporation of legacy information, the ease of use of the user interface, the security of the stored information, and the robustness of the software to support multiple file types and allow for the sharing of information across platforms.

Collaboration with Raytheon, a defense and aerospace systems company, allowed our team to develop and demonstrate a successful adoption of semantic abilities by a commercial company. Salient features of this work include a new tool, the e-Design *MemoExtractor Software Tool*, designed to mine and capture company information, a Raytheon-specific extension to the e-Design Framework, and a novel semantic environment in the form of a customized semantic wiki *SMW+*. The advantages of this approach are discussed in the context of the industrial case study with Raytheon.

1. INTRODUCTION

In recent years researchers have spent considerable effort developing semantic knowledge frameworks to capture, store, and facilitate the reuse of important design information [1-6]. The description logic offered by ontologies provides a means to add structure to often unstructured information [2, 7-10]. This additional structure has been used to improve querying capabilities of stored knowledge, therefore improving knowledge reuse [1, 9]. While much research has focused on the development of ontologies, very little has been done to implement the developed knowledge frameworks within an actual industry setting.

The research discussed in this paper addresses and provides approaches to overcome the difficulties associated with the adoption of ontological frameworks in industry. The research at UMass Amherst and the National Science Foundation Center for e-Design over the past five years has put our research team in a unique position to develop, implement, and evaluate such a transition. Our research has focused on the development of ontological frameworks to facilitate the structured representation of engineering modeling knowledge associated with conceptual design, engineering analysis, optimization, and decision making. This background provides us with a unique perspective on how ontological knowledge representation can be used to best enhance product development at companies and what steps are needed to allow for their implementation [3, 5, 6, 11].

This paper details the overall process utilized to enable the industrial implementation of semantic technologies, with specific sections addressing the various requirements of its adoption and the process used to meet them. Section 2 describes the related works and the significance of our presented work along with describing, comparing and contrasting the current semantic technology alternatives that are available for implementation. Section 3 outlines the industry needs and the implementation complications that were identified through our collaborative study with Raytheon, a defense and aerospace systems company (Sudbury, MA). Section 4 highlights the techniques employed in a collaborative case-study project with Raytheon along with the discussion of the results. Finally, the paper concludes with an overall discussion of the research and the steps necessary for the successful adoption of knowledge representation systems in industry.

2. BACKGROUND

2.1. Towards New Knowledge Management Systems

Many companies, both small and large, have felt it necessary to move beyond traditional design notebooks to digitally capture both product development and manufacturing related information. [12]. Therefore, the use of information technologies to model, store, and reuse knowledge in support of product development processes has become essential in today's industry. In general, the tools to meet these information needs are being developed from three different perspectives: 1) proprietary, specialized software by the company using the information [12], 2) as an extension of a current company software provider (e.g., PTC's Windchill [12, 13]), or 3) by a third-party as a piece of standalone software (e.g., Vistagy's EnCapta [14]).

Software developers often take disparate approaches when developing their knowledge management software, resulting in the use of proprietary methods to represent and share information [15]. When instantiating knowledge in a proprietary database, the developer can essentially build from the bottom up, elaborating and redefining on how knowledge is received, stored, and reused. While there are obvious benefits to this often closed-system approach, it becomes difficult to adapt architectures to suit different needs, such as adding or extending domains. These proprietary tools are rarely interoperable and require a significant amount of commitment for their implementation. Tools that support more open architectures, such as some emerging semantic technologies, are often more useful when addressing such needs.

Semantic techniques were utilized by McMahon *et al.* to develop one of these tools, their Waypoint system [16]. This system is intended for employment as an engineering document search and retrieval tool. Through automatic semantic classification of documents, the system enables search refinements in order to focus queries toward relevant and specific results. A case study implemented the system into the Airbus UK network. A large number of electronic documents were classified with taxonomies that were preexistent in the Airbus UK organization. Although the successfulness of the Waypoint system is still being determined, the semantic relationships as a 'back-end' and a continual feedback system as a 'front-end' user interface allows this system to have large potential to meet their specific goals. This system also has some limitations though. The Waypoint system relies on company documents initially being in a structured electronic form, as opposed to paper copies of documents. Further, even with the preexisting organization taxonomies, the system is not able to make inferences about the back-end ontology as many emerging semantic technologies allow.

To overcome the seeming lack of preexisting organization taxonomies and classifications, researchers at Purdue University have placed considerable interest in the acquisition of engineering information and the development of an ontology-based design document analysis and retrieval tool (ODART) [17-19]. This research addresses the need for a facilitated and more automated way to index and locate design documentation. This system mainly focused on the information retrieval and classification portion of the overall knowledge acquisition and reuse process, leaving other aspects of the process unaddressed. The system does not include the ability to allow users to interact with documents in a central location nor the ability to employ security restrictions on sensitive material. These types of system limitations have led to the growing use of Semantic Web technologies.

2.2. Semantic Web Technologies

An extension of the World Wide Web, the Semantic Web offers a both human readable and computable environment for capturing knowledge [20]. Though the Semantic Web is still evolving, it has become a popular means to share and store information in the development of knowledge bases. Semantic Web-driven ontologies and description logic have been used to provide formal standardized representations for engineering design knowledge. Additionally, the Semantic Web provides an environment where ontologies can be built with distributive and collaborative efforts.

While the use of ontology-based applications for knowledge management has yet to become mainstream in industry, the development of software applications based on semantic technologies is no longer in its infant stages. Many semantic applications are tailored towards improved knowledge management in distributed environments and improved knowledge retrieval. These applications range from locally run applications to large semantic knowledge repositories. In reviewing the available implementation methods, we first classified applications into three distinct categories: 1) local applications with local storage, 2) local applications with central storage, and 3) web-based applications with central storage.

The locally downloadable applications that store files in local repositories, such as *Protégé* [21] and *Swoop* [22], support the confined development and storage of semantic information. These tools have been mostly created as ontology development tools, however, and not as knowledge management tools. Downloadable applications with local repositories are an effective means for creating and developing ontologies but lack the appeal and versatility of more collaboration-oriented applications.

Another option is locally downloadable applications with central repositories, such as *pOWL* [23]. These applications provide a means to collaboratively develop ontologies and store them in a central knowledge base. *pOWL* in particular is marginally tailored towards knowledge management as it was developed as an easily deployable and easy-to-use, web-based ontology editing and publishing solution. While such tools provide a few more qualities desired of a knowledge management tool, i.e., a central repository, they require local installations which can be difficult to properly configure and maintain.

A third available option is the use of browser applications with central repositories. These applications appear to provide the most promise as a practical means for providing an industry end-user with a semantically enhanced knowledge management tool. Among browser applications with central repositories, the 'wiki' has become a popular method for the development of collaborative development [24]. A wiki is defined by the Merriam-Webster dictionary as "a Web site that allows visitors to make changes, contributions, or corrections" [25]. A recent wiki development, the 'semantic wiki' promises to improve upon current wiki abilities with semantic technologies [26]. Text-based semantic wikis offer the capabilities of a traditional wiki with semantic enhancements. These enhancements often include form-based data entry, back end triple-stores (Jena, Sesame), improved querying methods, and additional capabilities offered through the storage of semantic content [27-30]. Alternatively, form-based wikis have been built on more structured platforms and, as the name suggests, are form-based. Form-based semantic wikis have their semantic structure inherently built into their architecture, using ontology definition to define how information is both navigated and displayed [27, 31].

2.3. Semantic Wiki Alternatives

Several possible semantic wiki alternatives were considered including *SMW+*, *AceWiki*, *KiWi*, *Knoodl*, *Sweetwiki*, *Swirll*, and *OntoWiki* [32]. These platforms were compared and contrasted in order to determine which application might best suit industry needs. Based on these assessments, a single platform was identified as the most practical knowledge management application able to support implementation into industry. Our team decided to focus on open-source applications as they allow direct customization and could be implemented and evaluated at no cost. Consequently, various closed-source applications, such as *Knoodl* [29], were removed from initial consideration. Due to length requirements of the paper, the assessments discussed in this section are limited to select application options.

The *KiWi* Project is a text-based wiki supported by funding from the European Community's Seventh Framework Program, and therefore has not only a significant, but also a stable support community [28]. *KiWi* utilizes semantic reasoning and classifications in a web-based environment, allowing for collaborative efforts through one centralized database of knowledge. Use cases with Logica, a leading IT and business services company, and Sun Microsystems, a major contributor to open-source software, have shown samples of industry implementation work. The Logica project worked to employ *KiWi* as a centralized tool for project managers to access, write, and store project documents and knowledge. Due to the vast amount of different storage systems and knowledge management tools used by software developers, the Sun Microsystems project with *KiWi* aimed to provide a central, standard method for all developers to manage their material in a more effective means. Although there have been these limited efforts to implement *KiWi* into industrial practice, *KiWi* lacks desired commercially supported software and any form-like abilities.

At their current stage of development, form-based wikis, such as *OntoWiki* [31] (successor to *pOWL*), do not seem to have reached the necessary point for implementation as an industry-supporting application. While significant development has been shown, the form-based wikis seem to currently maintain a smaller support group and have less available extensions than their text-based counterparts. Although form-based semantic wikis are not currently considered a top alternative, they will be re-evaluated in the future in anticipation that their support and extensions will improve.

The most appropriate, currently available option was determined to be the semantic media wiki *SMW+*. *SMW+* is an open-source, text-based wiki supported by Ontoprise [30]. Positive attributes of *SMW+* not only include community development, but also commercial support, if necessary. *SMW+* also has a significant user support community and additionally offers multiple extensions to the basic semantic wiki. These extensions include the Halo extension, which offers support for semantic forms. With a semantic forms extension, it was expected that *SMW+* would give the text-based wiki the feel of a form-based wiki. *SMW+* uses the Jena reasoner as its back-end triple store, allowing for easy integration with a MySQL database. *SMW+* also provides a Microsoft Excel bridge which allows the query information generated in the wiki search to be imported into Excel for further calculations. Our research and the popularity of *SMW+* indicated a large amount of promise associated with the system. Ultimately, it was decided that the commercial readiness, commercial and community support, and greater versatility made *SMW+* the most appropriate application for a prototype example of industry implementation.

3. IDENTIFICATON OF INDUSTRY NEEDS

While academia has continued to make advancements in knowledge management through the use of ontologies and the Semantic Web, industry has been slow to follow. In an attempt to understand this perceived reluctance from industry, the UMass Amherst Center for e-Design has collaborated with Raytheon. Through this work our team has identified several obstacles in the form of unaddressed industry requirements which we believe have impeded industrial adoption of semantic technologies. This section defines five main industry requirements, as identified by our team, which should be present in the development of an industry-ready application.

These needs include the incorporation of key legacy information, the ease of new knowledge management software adoption, the robustness of the software to support multiple file types and allow for the sharing of information across platforms, the security of the stored information, and the ease of use of the user interface. A flowchart outlining the five main needs and their associated requirements can be found in Figure 1. The process is represented in a flowchart to represent the order in which our team chose to address the requirements. This flowchart was adopted as the foundation of an overall plan for the adoption of semantic techniques into industry.

3.1. Incorporation of Legacy Information

Much of engineering design consists of redesign. Due to this, the ability to incorporate legacy information is critical for supporting engineering design [16]. Without this essential ability, a company runs the risk of losing vital knowledge and information used in the development of their products. The loss of other critical information, such as company practices, also could prove harmful, leading to further importance for the comprehensive incorporation of legacy knowledge. This requirement involves the identification and mining of key categorical information in order to accurately associate the captured knowledge with associated instances.

3.2. Ease of Adoption

When a company adopts a new software application, it must take into consideration any possible challenges that may be encountered during its implementation and operation. These challenges may include installation and maintenance of the new software, new training requirements to operate and manage the software, training requirements to overcome the learning curve as an end-user to the software, compatibility of the new software with currently existing applications, and how the software will be installed to run with the existing infrastructure. In addition to these commonly encountered challenges, industry must also consider the requirements to adapt the knowledge framework to individual company needs. These more specific needs result from customized structure and relationships that relate to the specific practices of the company [17, 19].

3.3. Robustness of File-Type Support

The ability of the knowledge system to be adaptable enough to support many types of information is an important factor to consider in the selection process, especially for large companies. Similar to the variability of different computer aided design tools, many different files types are used to store information in engineering and industry [15]. The ways information can be generated and stored range from solid modeling software, to analysis software, to word processing software, to statistical analysis software, etc. Knowledge surrounding one individual product can include information stored in different file types from all of these, often proprietary, software formats. The ability to support multiple file types allows all product information to be stored in a single database, reducing redundancy throughout the entire database and improving knowledge storage and retrieval.

3.4. Security

For many companies and organizations, security of their sensitive information is one of the more, if not the most, important requirements for a database and knowledge management system [33]. Information security is necessary to not only keep competitive advantages but often to ensure required confidentiality. The security issues apply to concerns about outside hackers and often include concerns from within an organization or company as well. Tight control over information is a requirement of any knowledge management tool. Security can range from the allowance of only company members to view files to only allowing high management employees file access.

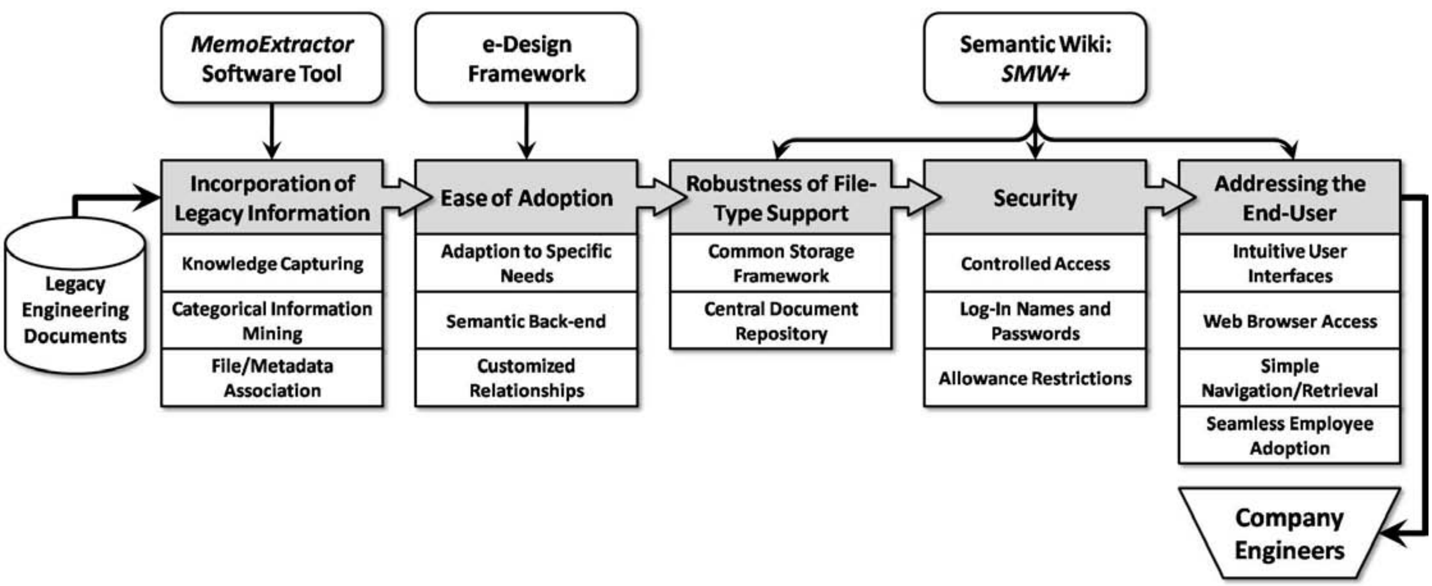

Figure 1 - Semantic Technique Adoption Flowchart

3.5. Addressing the End-User

There are many important concerns over end-user conformity within industry. Industry is constantly searching for means to streamline production and reduce costs. To this end employees are often asked to learn new tools in hopes of improving efficiency. Constant changes in company practices, however, can become discouraging and can cause employees to become reluctant to adopt new practices. This is an obstacle faced by any software that is introduced to industry [34]. Many employees may feel that it would be more efficient to continue to use previously learned methods as opposed to spending the time to familiarize themselves with a new method. The internal company inconsistencies caused by employees using multiple different software systems for the same process proves the importance of a universal, intuitive, familiar, and easy-to-adopt application.

3.6. Summary

Our team has identified five main requirements that we found must be addressed for a successful industrial implementation of semantic technologies. In addressing these requirements, we found it was necessary to not only enhance and extend existing technology, but to create and develop new technology as well. This next section will discuss how our group approached each of these requirements during the development of an industrial application.

4. SEMANTIC FRAMEWORK FOR INDUSTRIAL APPLICATION

In collaboration with Raytheon, we developed a semantic application for mining and managing information stored in both current and legacy engineering documents and files, including thousands of engineering memorandums, from a mechanical engineering directorate within the company. These documents and files contain important information on topics such as products, projects, and simulations. It was in Raytheon's interest to categorize, structure, and store these documents and files in a way that engineers could easily search and retrieve them. This section details and describes the challenges that were faced and the methods that were employed to overcome these issues during its development. The methods employed to address each of the five requirements can also be found in Figure 1.

Our first requirement, the incorporation of legacy information (3.1) was satisfied by the development of a Python-based tool named *MemoExtractor*. Details surrounding the development and use of this software tool are discussed in section 4.1. To address the need for ease of adoption (3.2), a customized extension of the e-Design Framework was created. The methods regarding the development of this semantic framework extension are detailed in section 4.2. The final three requirements, robustness of file-type support (3.3), security (3.4), and addressing the end-user (3.5), were overcome through customization of a semantic wiki, *SMW+*. Specific configurations of *SMW+* allow for the appropriate levels of file-type support along with an intuitive, familiar user interface. *SMW+* also allows for advanced levels of system security with its built in security features. The customization of this application is discussed further in section 4.3.

4.1. Development of *MemoExtractor*: A Tool for Knowledge Capturing

It was necessary to overcome the challenges associated with the incorporation of legacy information (3.1). Ontologies rely on the formal structuring of information, and converting currently unstructured information to the new structured format could result in the loss of vital information. This potential loss forced us to consider the issues of how to structure the information and what information needed to be structured. With companies already moving towards more computational databases, many different degrees of structure may be encountered.

Over 20,000 engineering memorandum documents regarding projects, analyses, products, and programs were identified as the legacy information that needed to be captured. Raytheon engineers desired a way to computationally utilize these paper-copy documents. The establishment of categorical information was necessary to individually instantiate and structure each memorandum document in a knowledge framework. The important categorical information was determined by Raytheon engineers to be the subject, author, recipient, date, memo number, and the project or program that the memorandum document relates to. This information was identified as the key information that further links related memos by their authors and more importantly, their project associations. When a Raytheon engineer needs information related to a certain project, these relations allow the system to present the engineer with only the properly associated documents.

Unlike other studies that were discussed in Section 2, the information that required capturing was not computationally available in electronic formats. To overcome this, it was necessary to develop a customized software application that had the specific ability to mine the required categorical information from the memorandum cover pages. Our research team developed the NSF Center for e-Design *MemoExtractor Software Tool* for this purpose. The structure of the *MemoExtractor Software Tool* process is shown in Figure 2.

After a preprocessing stage of scanning the memorandum documents into a flat PDF format and running them through a commercially available optical character recognition (OCR) process, the PDF files were exported to plain text files, which are associated to their respective PDF file by a simple tag. The Python[1]-based *MemoExtractor Software Tool* has the ability to read the generated plain text files, identify the locations of the categorical information provided by the framework knowledge structure, and then mine the relevant information. The tool then created and compiled an organized index of each memorandum file name, file location, and the associated categorical information.

[1] www.python.org

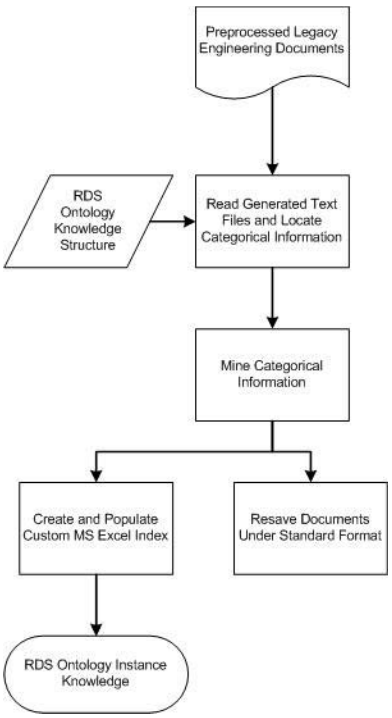

Figure 2 - MemoExtractor Software Tool Process Structure

Additionally, the *MemoExtractor Software Tool* rewrites the file names of the memorandum documents in a standard format. This allows further structure and organization to the file database. The mined information was later utilized in the customized ontology as instance knowledge, described in section 4.2, to ensure the correct properties were associated with each individual document instance. Although the *MemoExtractor Software Tool* was developed to mine custom information from specific Raytheon memorandum cover page formats, there is the ability to customize the application to fit other document types and information needs.

The *MemoExtractor Software Tool* was found to work consistently on a variety of over 12,000 Raytheon documents with an 80% overall success rate. The final number of processed documents will exceed 20,000. Success was determined by comparing the total number of information fields (subject, memo number, author, date, and recipient) that required fulfillment to the number of information fields that were actually filled. If the program was not successful in filling a field, the word 'None' was reported for that field.

The robustness of the tool was also displayed by the success of mining information from all variations of documents without fail. The variability of the documents included hand-written notes, graphs, figures, memorandums without consistent cover page templates, inconsistent quality of some native paper-copy memos before scanning, and memorandum documents dating back to the 1950's. The issues causing unsuccessful information mining were investigated as well. It was determined that failure to report information for a given field was caused by limitations with current OCR technology and limitations with PDF-to-TEXT file conversion. These issues go beyond the in-house development of the *MemoExtractor Software Tool* and therefore will not be addressed here.

4.2. Development of e-Design Framework Extension

Ontology-based applications can present a unique challenge in the realm of industrial adoption, as they often require someone with knowledge engineering experience. They also present unique advantages, however, as they can offer an adaptable framework which can be easily customized depending on the needs of the faction requiring the knowledge management application. Additionally, those based on the Semantic Web can provide a distributed environment with web accessibility, eliminating the need for local installations of a knowledge management tool.

Researchers at UMass Amherst and the National Science Foundation Center for e-Design have developed a comprehensive ontological framework in support of product development. The e-Design Framework is a complex network of semantic knowledge structures, including a decision making ontology, an engineering analysis model ontology, and an optimization ontology. The e-Design Framework allows for a large amount of customization and tailoring towards specific needs. This ability was essential to overcome the second challenge of facilitating the adoption of the semantic techniques into industrial practice (3.2).

After it was identified that Raytheon was interested in

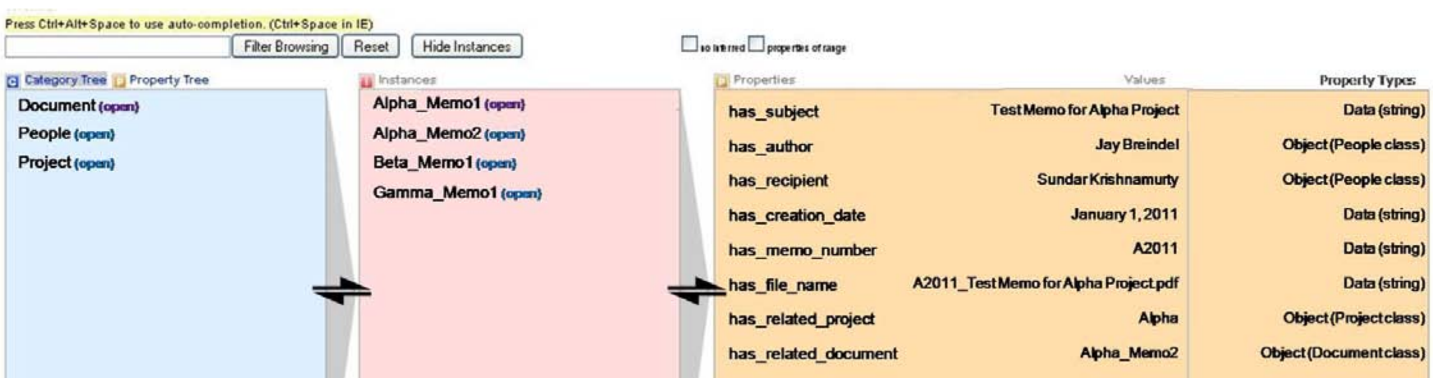

Figure 3 - Raytheon Document Support Ontology Semantic Structure in *SMW+* Ontology Browser

developing a structured database to store and retrieve important engineering memorandum documents, it was recognized that a customized framework was necessary. To address this issue, we developed a customized extension of the e-Design Framework that was specific to Raytheon's needs and practices. The Raytheon Document Support (RDS) extension was created in the ontology development tool *Protégé*.

In the RDS extension, each memorandum file was represented as an instance. The identification categories, discussed in Section 4.1, were created as properties linking and relating the memorandum document instances. The semantic structure of the RDS ontology extension is shown in Figure 3(application screenshot). The blue box depicts the three main classes of the ontology. These classes include a Document class, a People class, and a Project class. Each of these main classes contains a set of instances. In Figure 3, sample instances for the Document class are shown in the red box. Associated with each instance is a collection of four object-type and four data-type properties. The data-type properties are represented by strings, numbers, and alphanumerical text while the object-type properties are represented by instances of the three main classes. Figure 3 shows the list of properties associated with the Alpha_Memo1 instance. The categorical information identified earlier is accounted for by the properties has_subject, has_author, has_recipient, has_creation_date, has_memo_number and has_related_project. Furthermore, the information regarding the file name and related documents is also captured. Tailoring the semantic framework to Raytheon's needs facilitated its adoption into their Directorate practices.

By utilizing an extension of the e-Design Framework, Raytheon can also take advantage of the full knowledge management system. The resulting customization has the potential to increase Raytheon engineers' information storage and retrieval capabilities without requiring additional adoption techniques on their part.

4.3. Configuration of *SMW+* Platform

We extended the e-Design Framework with additional Semantic Web technologies to address the final three requirements, namely, robustness of file-type support (3.3), security (3.4), and addressing the end-user (3.5). Based on commercial readiness, commercial and community support, and greater versatility, *SMW+* was identified as the most appropriate application platform for the prototype example of industry implementation.

SMW+ was configured to include the semantic structure of the customized RDS ontology extension. Figure 3 shows the ontology browser available within *SMW+*. Through the custom configuration, each instance from the Document, People, and Project classes is created as an individual page. On each page, the properties associated with each instance are listed, including links to the pages designated for related projects and documents.

One of the advantages of ontologies is the allowance for a common framework to support multiple information types, such as floats, strings, and objects. While ontologies provide a significant amount of adaptability, the software application and language they are implemented in ultimately determine file type limitations and robustness capabilities of the knowledge management tool. *SMW+* allows for a large amount of possible file support. These file types include document files (i.e., .doc, .pps, .ppt, .xls, and .pdf), image files (i.e. .gif, .jpeg, and .png), audio files (i.e. .ac3, .mp3, and .ogg) and video files (i.e., .avi, .mov, .mpeg, and .wmv) [30]. These are only a select sample of the numerous file types that are supported by the implemented *SMW+* application. Specific configuration of *SMW+* enabled the system to properly account for the numerous PDF memorandum documents as well as allow for the possibility of including other file formats that may be added as the system progresses. The broad range of file support allows for one central database to connect all types of company information.

Although our team has conducted research regarding methods to control knowledge access by allocating permissions to portions of a semantic knowledge base, ontologies themselves do not currently offer any security advantages or disadvantages [35]. At this time this distinction is better suited for individual software applications. Certain applications allow for their own security methods and protocols, some suited better for industrial purposes than others. The selected application, *SMW+*, allows for access control where availability of sensitive and confidential content is restricted to certain employees and upper management [30]. The access is controlled through security protocols, log-in names and passwords, and separate allowances for each wiki piece of information.

The presentation of ontologies is often associated with such ontology development tools as *Protégé* and *Swoop*. While these tools provide a convenient and capable means for ontology development, they are not tailored to an untrained end-user. For industry to consider ontologies as an acceptable method to capture design knowledge, concerns must be addressed with user-friendly applications that do not require a familiarity with knowledge engineering or unfamiliar types of interfaces. An end-user application must be generally intuitive and familiar to allow for the smoothest possible transition from formerly used applications.

One of the most important aspects of a new application was to include an intuitive user interface to promote the seamless adoption and integration into an employee's everyday activities within Raytheon. The *SMW+* wiki skin, or interface, was custom chosen to present a familiar Wikipedia[2] type interface. This interface skin allowed one-click navigation between memorandum instances and utilized the semantic search features to allow user facilitated search and retrieval of desired documents. *SMW+* also allows for familiar graphic user interfaces for uploading and downloading documents to attach to each document instance. With this customized application, users are able to take advantage of interfaces that are similar to many other common software tools employed throughout industry.

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

5. DISCUSSION

As technology advances and knowledge capturing techniques improve, it is unfortunate that many semantic techniques are not being utilized by the companies they are designed to assist. Separate research has focused on the development and usefulness of semantic frameworks, the methods to extract the valuable information from company documents, and the technologies to interface with the semantic framework. Unfortunately, not enough has been done to incorporate all of these topics in one implementation method to manage company knowledge. Currently, there is a gap in technology between the creation of knowledge frameworks and their implementation, causing industry hesitance to use a foreign approach to manage their information. This research concentrated on the identification of important industry needs and the identification of potential methods to bridge this gap and utilize the knowledge frameworks that have already been developed.

Through this research, our team identified the five main industry requirements that we believe are necessary to address for a successful transition to semantic knowledge frameworks. These requirements included the ease of new knowledge management software adoption, the incorporation of legacy information, the ease of use of the user interface, the security of the stored information, and the robustness of the software to support multiple file types and allow for the sharing of information across platforms. Research was conducted on available semantic technologies that could contribute to this project. Many obtainable applications, such as semantic wikis, were extensively investigated. Their many benefits and drawbacks were presented and discussed. The research on these available applications was important due to the variance between specific company needs and processes. A company's individual needs may result in requirements for a different type, or multiple types, of semantic technology depending on the advantages certain applications provide. Although *SMW+* was successfully customized and configured as the semantic wiki application in our study with Raytheon, it may not be the best option for a different industry member.

To demonstrate how semantic applications can be effectively employed for useful purposes, we conducted an implementation case study with Raytheon. Through our collaboration, necessary information and overall structure were identified and mined from scanned engineering memorandum documents. This was accomplished through development and use of our *MemoExtractor Software Tool*. The mined information was then used to populate a customized Raytheon Document Support extension to the e-Design Framework, which gave company-specific instance knowledge to an *SMW+* application and intuitive user interface.

Overall, our identification of industry-related requirements, along with the development of methods to address these needs, provide a foundation for accurately and comprehensively capturing legacy knowledge within a semantic framework. Currently, steps are being taken to measure the overall success of the *SMW+* employment. These steps include measuring the amount of use of the system by Raytheon engineers, determining metrics to compare the effectiveness of the search and retrieval abilities, and gauging the overall employee satisfaction of interacting with the new front-end user interface. We expect that this research will lead to increased support and adoption of semantic technology within the industrial setting.

ACKNOWLEDGMENTS

This research was made possible by funding from the NSF I/UCRC Division of Industrial Innovation and Partnerships Award Number 0838747.

DISCLAIMER

Certain commercial products are identified in this paper. These products were used only for demonstration purposes. This use does not imply approval or endorsement by National Institute of Standards and Technology, not does it imply that these products are necessarily the best for the purpose.

REFERENCES

[1] S. Ahmed, S. Kim and K. M. Wallace, "A Methodology for Creating Ontologies for Engineering Design," *Journal of Computing and Information Science in Engineering*, vol. 7, pp. 132, 2007.

[2] B. Chandrasekaran, J. R. Josephson and V. R. Benjamins, "What Are Ontologies, and Why Do We Need Them?" *IEEE Intelligent Systems their Applications*, vol. 14, pp. 20, 2002.

[3] I. R. Grosse, J. M. Milton-Benoit and J. C. Wileden, "Ontologies for Supporting Engineering Analysis Models," *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, vol. 19, pp. 1-15, 2005.

[4] R. Mizoguchi and J. Bourdeau, "Using Ontological Engineering to Overcome Common AI-ED Problems," *International Journal of Artificial Intelligence in Education*, vol. 11, pp. 1, 2000.

[5] J. Rockwell, I. R. Grosse, S. Krishnamurty and J. C. Wileden, "A decision support ontology for collaborative decision making in engineering design," in *Collaborative Technologies and Systems, 2009. CTS '09. International Symposium* , pp. 1-9., 2009

[6] P. Witherell, S. Krishnamurty and I. R. Grosse, "Ontologies for Supporting Engineering Design Optimization," *Journal of Computing and Information Science in Engineering*, vol. 7, pp. 141, 2007.

[7] N. Noy and D. McGuinness, "Ontology Development 101: A Guide to Creating Your First Ontology," *Technical Report*, vol. KSL-01-05, 2001.

[8] B. Swartout, R. Patil, K. Knight and T. Russ, "Toward Distributed Use of Large-Scale Ontologies," *AAAI*, vol. SS-97-06, 1997.

[9] T. R. Gruber, "Toward Principles for the Design of Ontologies Used for Knowledge Sharing," *International Journal of Human-Computer Studies*, vol. 43, pp. 907, 1995.

[10] P. Zhan, U. Jayaram, S. Jayaram, O. Kim and L. Zhu, "Knowledge Representation and Ontology Mapping Methods for Product Data in Engineering Applications," *Journal of Computing and Information Science in Engineering,* vol. 10, pp. 021004, 2010.
[11] R. Fernandes, I. R. Grosse, S. Krishnamurty and J. C. Wileden, "Design and Innovative Methodologies in a Semantic Framework," *ASME Conference Proceedings,* pp. 237, 2007.
[12] D. T. Liu and X. W. Xu, "A Review of Web-Based Product Data Management Systems," *Comput. Ind.,* vol. 44, pp. 251, 2001.
[13] PTC, "PTC Products: Windchill," 2011. http://www.ptc.com/products/windchill/.
[14] Vistagy, "Enable Innovation with Specialized Engineering Software Powered by EnCapta," 2011. http://www.vistagy.com/products/index.aspx.
[15] J. Altidor, J. C. Wileden, Y. Wang, L. Hanayneh and Y. Wang, "Analyzing and Implementing a Feature Mapping Approach to CAD System Interoperability," *ASME Conference Proceedings,* 2009.
[16] C. McMahon, A. Lowe, S. Culley, M. Corderoy, R. Crossland, T. Shah and D. Stewart, "Waypoint: An Integrated Search and Retrieval System for Engineering Documents," *Journal of Computing and Information Science in Engineering,* vol. 4, pp. 329-342, 2004.
[17] Z. Li and K. Ramani, "Ontology-based Design Information Extraction and Retrieval," *Artificial Intelligence for Engineering Design, Analysis and Manufacturing,* vol. 21, pp. 137-147, 2007.
[18] Z. Li, V. Raskin and K. Ramani, "Developing Engineering Ontology for Information Retrieval," *Journal of Computing and Information Science in Engineering,* vol. 8, pp. 011003, 2008.
[19] Z. Li, M. C. Yang and K. Ramani, "A Methodology for Engineering Ontology Acquisition and Validation," *Artificial Intelligence for Engineering Design, Analysis and Manufacturing,* vol. 23-31, pp. 37, 2009.
[20] T. Berners-Lee, J. Hendler and O. Lassila, "The Semantic Web," *Sci. Am.,* vol. 284, pp. 34, 2001.
[21] Stanford University "Protégé," 2010 http://protege.stanford.edu/.
[22] A. Kalyanpur, B. Parsia, E. Sirin, B. C. Grau and J. Hendler, "Swoop: A Web Ontology Editing Browser," *Web Semantics: Science, Services and Agents on the World Wide Web,* vol. 4, pp. 144-151, 2006.
[23] S. Auer, "pOWL - A Web Based Platform for Collaborative Semantic Web Development," http://powl.sourceforge.net/overview.php.
[24] B. Leuf and W. Cunningham, *The Wiki Way: Collaboration and Sharing on the Internet.* Addison-Wesley Professional, 2001.
[25] Merriam-Webster Dictionary, "Wiki Definition," 2011. http://www.merriam-webster.com/dictionary/wiki.
[26] D. Fowler, R. M. Crowder, T. Guan, N. Shadbolt and G. Wills, "Requirements for semantic web applications in engineering," in Montreal, Quebec, Canada, pp. 1-11, 2010.
[27] M. Krötzsch, S. Schaffert and D. Vrandecic, *Reasoning in Semantic Wikis.* 2007.
[28] KiWi, "KiWi Project Home Page," http://www.kiwi-project.eu/.
[29] I. Revelytix, "Knoodl Home Page," 2009. http://knoodl.com/ui/home.html.
[30] Ontoprise, "SMW+: Semantic Enterprise Wiki Homepage," 2010. http://wiki.ontoprise.de/smwforum/index.php/Main_Page.
[31] S. Dietzold, "Ontowiki Home Page," 2010. http://ontowiki.net/Projects/OntoWiki.
[32] Semantic Web, "Semantic Wiki Projects," February 11, 2011. http://semanticweb.org/wiki/Semantic_wiki_projects.
[33] A. M. Eskicioglu, "Protecting Intellectual Property in Digital Multimedia Networks," *Journal Computer,* vol. 36, pp. 39, July, 2003.
[34] A. Abran, A. Khelifi, W. Suryn and A. Seffah, "Usability Meanings and Interpretations in ISO Standards," *Software Quality Journal,* vol. 11, pp. 325-338, 2003.
[35] N. Kanuri, I. R. Grosse, J. C. Wileden and W. -. Chiang, "Ontologies and fine-grained control over sharing of engineering modeling knowledge in a web based engineering environment," pp. 105-113, 2005.