ENRICHING PRODUCT INFORMATION DURING THE PRODUCT LIFECYCLE

Kary Främling *, Lutz Rabe **

**) Helsinki University of Technology, P.O.Box 5500, FI-02015 TKK, Finland*
***) BIBA, Hochschulring 20, 28359 Bremen, Germany*

Abstract: A manufactured product usually has a great amount of design knowledge and other product information associated with it in CAD drawings and other technical documentation. Today, such product information mainly enables preventive maintenance of products. By collecting usage, performance and maintenance information about the product, it could be possible to perform more predictive maintenance. The key issue is to document product information in a way that would enable analysing the relation between usage parameters and actual product failures. Semantic networks are here proposed as an intermediate knowledge representation format for combining such static and dynamic product information. *Copyright © 2006 IFAC*

Keywords: knowledge representation, semantic networks, intelligent machines, decision support system.

1. INTRODUCTION

When studying available commercial Product Lifecycle Management (PLM) solutions, it appears that they fulfil most needs for managing information about product types, including the management of versions and product variants. If we change the definition of PLM from managing product types into managing the lifecycle of product items, then current PLM solutions generally only cover the beginning of life (BOL) phase of the lifecycle, i.e design and manufacturing. Furthermore, the functionality of existing PLM systems mainly seems to be guaranteed in the context of one single company or a group of tightly collaborating companies.

For individual product items, the middle of life (MOL) and end of life (EOL) phases are becoming increasingly important to manage. A main MOL issue is to keep the product usable in a way that is as close to optimal as possible through efficient maintenance procedures. When this is no longer possible, the product item should be disposed of or recycled in a way that minimizes waste of material and energy. Managing the MOL and EOL phases efficiently is essentially an information management task. When failures are detected and corrected during maintenance of the product item, it would be essential to document the failures and corrective actions as well as possible. Otherwise they remain tacit knowledge of the maintenance engineer, which tends to become lost e.g. if the maintenance engineer leaves the company for some reason. In this paper, we attempt to point out technologies that could be used for documenting this tacit knowledge in such a way that it would be preserved and usable in the future as easily as possible.

However, just documenting situations of failure, maintenance and repair is not alone sufficient for guaranteeing the optimal operation of product items. In order to avoid product failures and breakdowns, predictive or condition-based maintenance should be the objective. Especially for condition-based maintenance it is necessary to associate occurring failures with changes in sensor signals or other indicators before the failure occurs. Such associations can be detected by different kinds of data analysis or machine learning methods. Such methods require that all analysed data be represented in formats that the method can handle. This is usually not a problem for sensor data, which can be handled by most formal methods. It is more challenging to transform fault reports of maintenance engineers into a format that can be handled by analysis methods.

In this paper we propose using general semantic networks as a “best compromise” for representing

product information. Semantic networks can represent semantic relations between different concepts used by designers, maintenance engineers and other human actors involved. By using connection strengths between nodes, semantic networks can be used for diagnostics and prognostics purposes. Connection strengths can be entered manually or as results from data analysis methods. Data analysis methods also exist that construct semantic networks themselves as a result of the analysis. A real-world application scenario on truck maintenance illustrates the use of semantic networks for condition-based maintenance.

The structure of the paper is as follows: after this introduction, section 2 gives an overview of the state-of-the-art in managing product-related information. Section 3 gives an introduction to semantic networks and attempts to show how both tacit knowledge and sensor data can be transformed into a suitable format for semantic networks, followed by conclusions.

2. BACKGROUND

Product information is usually classified as being *data*, *information* or *knowledge* (and possibly even *understanding* and *wisdom*). Unfortunately there is no universally accepted and exact definition of these terms, even though a consensus exists about their general meaning. Data is “raw”, which means that it has no meaning of itself, not even units or other associated semantics. When adding semantics to data, e.g. units and source where the data comes from, the data becomes information. Knowledge is the ability to use information in some “advised” way, e.g. answering “how” questions or producing new information based on existing information. In the rest of this paper, data, information and knowledge will be used according to these guidelines. When speaking about all the three together, we will use the term “*product information*”.

2.1 Knowledge management techniques for externalisation of tacit knowledge

Knowledge Management (KM) can commonly be considered as the collection of processes that govern the creation, dissemination, and leveraging of knowledge to fulfil organisational objectives. Although this definition is useful for expressing the overall spirit of KM, it seems to be insufficient in the context of this work. This is mainly because management of product information in the context of this paper intends to cover the complete lifecycle of a product. Thus, KM processes are often of an inter-organisational nature between various actors that are rather loosely coupled over the time. Consequently KM has to fulfil more than one partner’s organisational objectives.

Closing the information loop between the processes of a product’s lifecycle for the dissemination of product information requires co-operation or collaboration between companies, organisations and individual parties such as product owners or freelancers. Thus, one major issue of KM between various parties is the co-operation between people who are acting under specific social, organisational, economic, strategic and legal conditions. Looking at inter-organisational co-operations from this perspective puts the individual employee and the knowledge that is bounded to him in the centre of the observations.

Nowadays there is a widespread agreement that considering knowledge requires the differentiation of knowledge that is already expressed or codified and knowledge that resides in the head of an individual. Along with this differentiation the concepts explicit and tacit knowledge were developed in the scientific community. Explicit knowledge, as the first word in the term implies, is knowledge that has been articulated and, more often than not, captured in the form of text, tables, diagrams, product specifications and so on. In a well-known and frequently cited 1991 Harvard Business Review article titled "The Knowledge Creating Company," Ikujiro Nonaka refers to explicit knowledge as "formal and systematic" and offers product specifications, scientific formulas and computer programs as examples. It can be expressed in words and numbers and is easily communicated and shared as hard data, scientific formulae, codified procedures or universal principles.

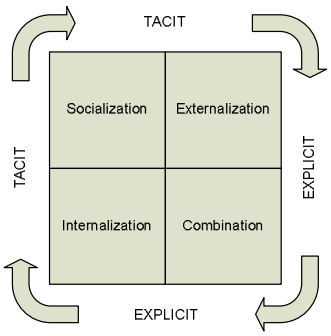

Fig. 1: Knowledge Spiral Model (Nonaka and Takeuchi, 1995)

Tacit knowledge is a phrase apparently coined by Michael Polanyi with his book Personal Knowledge (Polanyi, 1962). He described tacit knowledge as a characteristic of experts, who often have no explicit theory of their work: they simply perform skilfully. For the following, tacit knowledge shall be defined as non-explicit knowledge, which basically means that tacit knowledge resides in the mind of a human. Tacit knowledge has been found to be a crucial input to the processes covered by a product’s lifecycle. Nonaka and Takeuchi (1995) see the mobilization and conversion of tacit knowledge as the key to knowledge creation. The difficulty inherent in the management of tacit knowledge is that experts and stakeholders of the product may not be aware and hence, unable to articulate, communicate and describe

what they know. With their SECI Model (Fig. 1), Nonaka and Takeuchi provide a structured base for approaching this problem. The interaction between tacit and explicit knowledge from lower to higher levels (individual, group, organizational, inter-organizational) emerges as a knowledge spiral that includes four layers of knowledge conversion: socialization (from tacit to tacit), externalization (from tacit to explicit), combination (from explicit to explicit), and internalization (from explicit to tacit).

A successful knowledge management application needs to solve the problem of capturing and exploiting tacit knowledge. Current solutions that focus on structured data and semi-structured documents are limited to the knowledge that users are willing or able to enter in these forms (Brown et al, 2001). Considering the process of knowledge externalization, where tacit knowledge needs to be articulated into explicit knowledge, two major issues are significant. The first one is the effort (in correlation with his current situation), which is required for an individual to articulate his knowledge; the second one is the expressiveness of the resulting explicit knowledge. Thus, individual barriers for the externalization of tacit knowledge can be reduced among others by decreasing the effort for the externalization or increasing the individual benefit.The remainder of this section will mainly address methodologies and technologies that support the knowledge conversation and transportation with particular respect to the aim of decreasing the individual barriers. Concerning the individual conversation processes based on the SECI model the proposed concept can be described as follows:

Socialisation: In the context of this paper, the socialization process requires the provision of communication means that support the inter-organizational speech based communication of the product's stakeholders. Communication means should be easy to use in order to decrease technical barriers in stakeholders' communication. Communication tools to be applied should consequently be based on standard office applications and/or Internet- or browser-based applications.

Externalization: Expressing or articulating knowledge requires basically a certain degree of endeavor for the involved individuals. Together with the missing benefit of this process for the participating individuals, this is the most restricting barrier in capturing tacit knowledge. One approach for reducing the effort of externalizing tacit knowledge is to couple capturing processes with already existing communication processes. A substantial amount of individual knowledge is conveyed by speech either in meetings, on the telephone, or in casual conversations. For example a phone conversation between the product owner and a customer service in case of product failure contains a lot of detailed information on the specific product, its particular conditions of usage etc. This conversation can be captured and stored as text files that initially contain rather unstructured data. Knowledge extraction techniques could then provide access to the underlying knowledge that an organization or a team wants to preserve. Thus, the externalization process consists of firstly capturing the spoken word into textual information and then applying categorization methodologies to increase the reusability of this data.

One approach to foster the externalization processes is to create or increase the benefit of the externalization processes for the individuals. In the example of a product failure given above, the user of the product might approach the service or even other users of the product by describing the details of the failure in a forum, a blog or a wiki to get a solution for his problem. A wiki is basically a group of Web pages that allows users to add content, as on an Internet forum, but also permits others (often completely unrestricted) to edit the content. In essence, the wiki is merely a vast simplification of the process of creating HTML pages, and thus is a very effective way to exchange information for solving problems through collaborative effort.

Combination and Internalization: To support the combination processes it is required to apply methods and techniques for the enrichment of rather unstructured data resulting from the socialization and externalization processes. To increase the degree of reuse of externalized knowledge that results from the voice recognition process, methodologies for text categorization (TC) and information retrieval (IR) can be applied. Results of the TC and IR processes can be improved by the application of semantic networks as shown by Gonzalo (Gonzalo et al, 1998) or Jiménez (Jiménez et al, 2003). Consequently TC and IR processes have to be based on semantic descriptions of product related information objects, i.e. their structure and relationships.

2.2 Static and dynamic product information

In the area of post-sales maintenance, product information is often separated into static product information and dynamic product information (Simon, *et al.*, 2001). This classification of product information also seems suitable for PLM. Static product information means the specification of the product type as manufactured, e.g. materials, components, used suppliers, configuration and options, servicing instructions and end-of-life information. Dynamic product information comprises information that is collected during the use of a product. This includes patterns of use, environmental conditions, servicing actions, and part replacements.

Different applications have their own characteristics in how product information is created, gathered, maintained and provided to the ones needing it. Marsh and Finch introduce the concepts of centralized and localized (or decentralized) storage

(Marsh and Finch, 1998). The localized solution involves storing product information on the component itself. In the centralized model, product information is transmitted and collected to one place, so it is easy to perform data analysis on entire product populations and detect general rules that can be used for improving product design and perform predictive maintenance, for instance. If dynamic product information on the product item level can be continually transmitted to the centralised system, then the centralised system can also be used for real-time monitoring of every product item. In practice, communication of dynamic product information is often intermittent so real-time processing of it becomes impossible with a purely centralised solution.

It should also be pointed out that a centralized solution does not necessarily mean that product information would be located in only one place. As enterprises become increasingly global and networked (the "virtual enterprise"), product information tends to become spread on computer systems of multiple companies. This means that accessing product information may require communication between information systems of many different organizations (Koç and Lee, 2001).

3. MANAGING PRODUCT INFORMATION REPRESENTED BY SEMANTIC NETWORKS IN A MULTI-ORGANISATIONAL SETTING

The need to combine static product information with dynamic product information can be illustrated by a real-world vehicle predictive maintenance scenario. When a new vehicle type is launched, maintenance is mainly preventive based on estimates and assumptions made during design and manufacturing, i.e. static product information. When a sufficient number of vehicles have been in use for sufficiently long, collected dynamic product information can be used for improving initial maintenance schedules. Dynamic product information consists of collected statistics, sensor values etc. from the vehicles. However, in order to use this data for creating or improving predictive maintenance, it is also necessary to transform fault reports and other tacit knowledge of maintenance engineers into a format that can be handled by analysis methods. Finally, resulting decision support models from the analysis methods for predictive maintenance should preferably be embedded in the vehicles themselves, whereby needs for maintenance could be detected in real-time.

The first procedure of translating tacit information into explicit information happens during the design phase where the designer's skills are translated into a product design and related documentation. CAD models produced during the design phase are one of the most essential parts of product information. CAD models are a form of semantic networks that explicitly model "part-of", "depends-on" and similar relationships. The bill-of-materials (BOM) used during manufacturing is another important piece of product information that can be represented as a semantic network. However, the BOM is usually representing product information on a product-type or product-variant level rather than on a product-item level. Therefore the classical BOM is not sufficient for managing dynamic product information, where parts of the product item have individual properties and where parts can be changed during the product lifecycle (Främling, *et al.*, 2006).

3.1 Semantic networks

Semantic networks represent sets of named relationships between different nodes (or objects) in a network. Using a collection of pair-wise relations between nodes, where every relation may also have an associated "strength", can represent a semantic network. Relation strengths are particularly useful when semantic networks are used for reasoning, e.g. for diagnostic or prognostic purposes as those needed in many MOL applications.

Solutions for managing semantic networks in a multi-organisational context are being developed under the name "semantic web". The Resource Description Framework (RDF) and Web Ontology Language (OWL) are examples of standards being developed for the semantic web. Software frameworks also exist that use these standards, e.g. Jena (http://jena.sourceforge.net/), OpenRDF (http://openrdf.org/) and the 7.2.3 Redland RDF Application Framework (http://librdf.org/).

However, RDF and OWL are mainly focused on describing web content rather then on describing product information. Furthermore, the related software tools are not, as such, designed to be used for implementing distributed applications. Therefore agent frameworks could be more suitable for this purpose. Examples of such agent frameworks are, ABLE (http://www.alphaworks.ibm.com/tech/able) and JADE (http://jade.tilab.com/) that integrate inter-organizational communication. In a multi-agent framework, agent references correspond to links between nodes of a semantic network. Therefore agent frameworks could be used as building blocks for a distributed implementation of semantic networks for describing product information.

When using an agent framework with support for data analysis and decision support, the nodes themselves can also be "intelligent". Especially the data analysis methods included in the ABLE framework could be applicable as decision support systems. ABLE data analysis and decision support agents provide support for many different data analysis methods, e.g. naïve Bayes, decision trees and neural networks. In addition to these, ABLE agents exist for both crisp and fuzzy rules that are useful for explicitly expressing expert knowledge. This portfolio of methods is particularly

interesting for MOL scenarios that include diagnostics, prognostics and condition-based maintenance. ABLE agents can be trained both on- and offline and included in different software components to perform filtering or decision-making on different levels.

3.2 Management of product information in a multi-organisational context

As long as the KM processes required for PLM can be performed in a centralized manner it is relatively easy because decisions on data formats and information processing methods can be taken locally. When PLM is distributed, standards become essential both concerning the format of the data and how to exchange it. The chronologically first possibility of conveying product information in electronic format was using EDI messages or some file-transfer protocol (Angeles and Nath, 2001). The logic of using EDI messages to convey product information is similar to paper-based transfer, the information is sent to the next downstream supply chain partner (i.e. the primary customer). Handling product information with electronic or paper-based messaging is problematic e.g. because of cost and unnecessary duplication of information (Främling, *et al.*, 2006):

The most popular way of overcoming these problems has been the use of Internet and Extranet technologies for transferring product information. In practice, forerunner companies have developed portal applications through which their customers can download the product information. Portals are Internet based services that can represent information from various systems in a single place, through a browser (Linthicum, 2001). Portals have proved to be efficient in making the product information available to supply chain participants and have thus grown increasingly common. However, portals are just the user interface representing the information to the consumer of the information. The issues related to the management of information, e.g. linking information to specific product items and handling the bill-of-materials of products, demand additional solutions.

Web Services (WS, see http://www.w3.org/2002/ws/) seem to be the most universally accepted concept for XML-based communication between software components or applications over the Internet. The WS activities include several protocols for different tasks, notably the Simple Object Access Protocol (SOAP, http://www.w3.org/TR/soap/) for XML-based messaging and the Web Services Description Language (WSDL, http://www.w3.org/TR/wsdl) for describing services and their interfaces. Especially SOAP and WSDL seem to gain in popularity for defining application interfaces. The main advantage of WS is that they are relatively easy to configure (especially compared to EDI-based solutions) and could even allow for dynamic discovery of appropriate services for a given task. Other XML-based communication protocols are ebXML (http://www.ebxml.org/) and RosettaNet (http://www.rosettanet.org/). WS are also the cornerstone of the EPC-IS (EPC Information Services, see http://www.epcglobalus.org /Network/EPC_Information_Services.html) under development. EPC-IS could become relevant for PLM but is not yet a ratified standard. If we assume that applications can communicate over the Internet, the next step would be to react to incoming new information. This could signify the filtering of dynamic field data, detecting anomalies, making decisions on how to react on anomalies etc.

3.3 Condition-based maintenance in a real-world truck scenario

Fig. 2 shows a simple semantic network that addresses how different nodes of the net affect the condition-based decision on whether the engine oil should be changed or not at the next garage stop of a truck. When using an agent framework, some nodes of the network would typically be agents, e.g. the neural net in Fig. 2. The estimation of the number of kilometres could also be done by other methods than a neural net and the nodes of the net could also be semantic networks themselves. Finally, relation names tend to be subjective unless a standardized ontology has been defined for the domain.

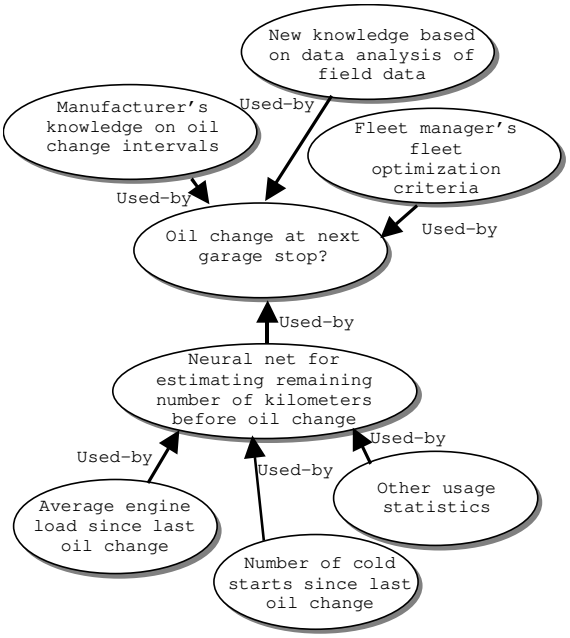

Fig. 2. Example of semantic network for predictive decision making on when to change engine oil.

In Fig. 2, explicit knowledge based on incoming field data is represented by the upward arrows towards the decision to make, i.e. should the oil be changed at the next stop or not. The downward arrows represent both knowledge that originally may have been tacit (manufacturer's knowledge on oil change intervals,

fleet manager's fleet optimisation) and knowledge obtained by data analysis of field data. In a scenario like this, a successful codification of both tacit and explicit knowledge is required in order to be able to take the appropriate decision. One important aspect of this scenario is also that several different "organisations" are involved, i.e. at least the truck, the manager of the truck fleet and the manufacturer.

The implementation of this scenario so far consists of three main components:

1. Universal Plug-n-Play (UPnP) enabled software that communicates directly with on-board diary for retrieving vehicle information.
2. Middleware software for inter-organizational communication (the DIALOG platform, http://dialog.hut.fi) (Främling, *et al.*, 2006).
3. Analysis agent(s) using the ABLE framework.

The DIALOG platform can use several different communication protocols for inter-organizational communication, notably WS protocols. It has support for basic operations such as tracking, retrieving product information and managing semantic relations between product items. It is also possible to register new agents that send and receive standard or application-specific messages. So far, methods for converting tacit to explicit knowledge have not been used in this scenario.

In the scenario of Fig. 2 the inter-organisational borders may in principle be located at any arrow. The DIALOG middleware can be used for communication at these borders. Vehicle information can be sent by an application-specific agent in the vehicle using DIALOG and be received by an identical agent at the manufacturer's end, for instance. In this basic scenario, the manufacturer's agent would then have access to both static and dynamic product information needed for taking the decision on whether an oil change is needed. If parts of the decision need to be made locally in the vehicle or by the fleet manager, more agents might need to be involved.

4. CONCLUSIONS

The initial assumption of this paper was that MOL and EOL processes such as maintenance or repair have to be supported by item-specific management of product information that covers the entire lifecycle of the product. Based on this assumption, this paper discussed the relevance of semantic networks as enabling technology in the context of a practical example. Based on the requirements of a real-world truck scenario we presented approaches for transforming existing tacit knowledge to explicit knowledge and for managing explicit knowledge in a multi-organisational context to support the MOL process of condition-based maintenance.

It is suggested that semantic networks would be applicable as a general tool for connecting static and dynamic product information for product lifecycle support. Specific semantic networks can improve the results of text categorisation and information retrieval processes required for the automatic inclusion of tacit knowledge to the overall MOL and EOL processes but also be used for diagnostics and prognostics required for condition-based maintenance processes. Although existing research in this area indicates the relevance of semantic networks in the presented approaches, further work is required to explore the particular requirements in integrating tacit knowledge in decision making processes and the role of semantic models as a connecting element for it.

REFERENCES

Angeles, R. and R. Nath (2001). Partner congruence in electronic data interchange (EDI)-enabled relationships. *Journal of Business Logistics*, **22:2**, 109-128.

Brown, E., S. Srinivasan, A. Coden, D. Ponceleor, J. Cooper, A. Ami and J. Piepe (2001). Towards Speech as a Knowledge Resource. *IBM Systems Journal*, **40:4**.

Främling, K., T. Ala-Risku, M. Kärkkäinen and J. Holmström (2006). Agent-based Model for Managing Composite Product Information. *Computers in Industry,* **57:1**.

Gonzalo, J., F. Verdejo, I. Chugur and J. Cigarrán (1998). Indexing withWordNet Synsets can improve Text Retrieval. In: *Proc. of the Workshop on Usage of WordNet for NLP*.

Jiménez, D., E. Ferretti, V. Vidal, P. Rosso, and C. F. Enguix (2003). The Influence of Semantics in IR using LSI and K-Means Clustering Techniques. In: Proc. of Workshop on Conceptual Information Retrieval and Clustering of Documents, 286–291. ACM Int. Conf., Dublin, Ireland.

Koç, M. and J. Lee (2001). A System Framework for Next-Generation E-Maintenance System, EcoDesign. In: *proc. of Second International Symposium on Environmentally Conscious Design and Inverse Manufacturing*. Tokyo Big Sight, Japan, Dec. 11–15, 2001.

Linthicum, D.S. (2001). *B2B application integration: e-business-enable your enterprise*. Addison-Wesley, Boston.

Marsh, L. and E. Finch (1998). Using portable data files in facilities management. *Facilities*, **16:1-2**, 21-28.

Nonaka, I. and H. Takeuchi (1995). *The knowledge creating company. How Japanese companies create dynamics of innovation*. Oxford University Press, New York, USA.

Polanyi, M. (1962). *Personal knowledge. Towards a postcritical philosophy*. University of Chicago Press, Chicago, USA.

Simon, M., G. Bee, P. Moore, J.S. Pu and C. Xie (2001). Modelling of the life cycle of products with data acquisition features. *Computers in Industry*, **45**, 111-122.